

**Igor Svoboda**

PhD Student, Department of Information Security

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

ORCID: 0009-0001-1709-2904

i.svoboda@kpi.ua

**METHOD FOR SYNTHESIZING AND PRIORITIZING CYBER INCIDENT SCENARIOS IN
CRITICAL INFRASTRUCTURE USING LLM AND AHP**

Abstract. The paper deals with the problem of linking each stage of a multi-stage cyberattack on critical infrastructure to the assets, accounts and operational control systems involved. Existing methods based on attack trees require a predefined system model, while approaches based on large language models (LLMs) do not check the organizational support of each generated attack step. This study aims to develop and experimentally evaluate an LLM-assisted decision-support method for the automated synthesis and multi-criteria prioritization of cyber incident scenarios. The proposed method is based on asymmetric generation: the organization in its initial operational state is decomposed hierarchically into assets, and causal event sequences are generated backward from a given incident. Structures generated by an LLM are validated against a JSON schema and checked for graph consistency, while the analyst identifies missing elements for local refinement. Complete scenarios are prioritized using the analytic hierarchy process (AHP). An event sequence is accepted only when every supporting object belongs to the organization’s hierarchy. Experimental evaluation used a modeled case of unauthorized switching that de-energizes substations of a regional electricity distribution operator in Ukraine. Four rounds of refinement recovered all events of the reference model with their causal relationships and object-event connections and produced five complete variants of attack scenarios. Four of these variants depended on a single relationship between the shared control service and the organizational hierarchy. Personnel credential compromise ranks first in the prioritization with a score of 0.816, ahead of the shorter dispatcher session takeover, and retains first place in all runs of the sensitivity analysis. These experimental results show that combining LLMs with formal graph checks detects attack steps that lack supporting organizational objects. In practice, security operations centers (SOCs) can use the approach to find critical dependencies of attacks and to decide which credentials and operational procedures to protect first.

Keywords: large language models; cyber incident scenarios; object-event synthesis; analytic hierarchy process; critical infrastructure; electricity distribution; decision support.

INTRODUCTION

The problem statement. The study is relevant because electricity distribution operators need to assess multi-stage cyberattacks that affect interconnected organizational and technical components. Each stage of these attacks involves particular accounts, services and control systems. Security planning for such attacks therefore requires methods that transform a specified target incident and a description of the organization into scenarios that specialists can examine stage by stage.

This problem is especially important for the Ukrainian energy sector, which accounted for 279 of the 5,927 cyber incidents processed in 2025 by CERT-UA, the Government Computer Emergency Response Team of Ukraine [1]. In the December 2015 attacks, the adversary reached operational systems through the compromised enterprise network, switched equipment without authorization and took control away from operators [2], [3]. The attack that was prevented in April 2022 targeted substations, Windows and Linux systems, as well as network equipment simultaneously [4]. Scenario analysis should therefore connect adversary actions across several technical and organizational areas.

Analysis of recent studies and publications. Existing methods describe how an incident develops from different starting points: fault tree analysis traces causes backward from an unwanted top event, while event tree analysis follows the consequences of an initiating event. Bow-tie models combine both analyses in one diagram [5]. Attack trees, which describe the routes to the attacker’s goal, preserve the temporal ordering of attack steps when sequential conjunction is used [6]. In [3], sequential decomposition is applied to known cyberattacks on control systems of critical infrastructure. However, these methods rely on a predefined description of the system under study.



Automated formalization of the organizational structures that underlie generated event sequences remains an important unresolved problem. Object-Process Methodology links persistent objects with the processes that change them [7]. Attack trees are generated from templates [8], combined safety-security models [9] and system specifications analyzed with artificial intelligence and natural language processing [10]. In recent years, large language models (LLMs) have extended automated model construction to less formal sources. Study [11] uses an LLM to chain known vulnerabilities based on their preconditions and effects and to turn threat reports into attack graphs. In [12], several LLM agents combine operational data with system design information to create attack patterns for an industrial control system testbed. Such patterns exceed those of human experts in number and quality. LLMs can also construct explicit world models for planning [13]. At the same time, these approaches require a predefined system model or do not check whether generated events are supported by organizational objects.

Work [14] applies an LLM guided by a reusable instruction template to convert incident descriptions into structured report records validated against a predefined data model. A retrieval-augmented generation (RAG) pipeline provides the requirements of regulatory documents. This approach reduces the time needed to prepare an incident report from hours to minutes but processes incidents that have already occurred. This study deals with an earlier stage, scenario synthesis. Such synthesis proceeds from two ends: it decomposes the organization in the initial state into objects and generates events backward from the target incident. Building the two structures in different directions makes the synthesis asymmetric.

The analytic hierarchy process (AHP) compares alternatives against several criteria [15]. Study [16] uses AHP in a decision support system for selecting cybersecurity architectures for critical infrastructure. In this system, the weights of the criteria, including detection and response times, automation level, cost and policy compliance, are updated as threats change. LLMs acting in assigned roles have also been used to weight transitions between the nodes of a corporate network [17] or to propose AHP criteria and alternatives with consistent pairwise judgments [18]. In [19], LLMs provide the judgments for a predefined analytic network process (ANP) model of security controls, criteria, threat vectors and constraints. The alternatives in these studies are security controls or paths in a given network. This paper, however, uses multi-step scenarios that are constructed together with their supporting organizational objects and verified before ranking.

Thus, the remaining gap is the synthesis of attack scenarios with verified connections between events and persistent organizational objects.

The paper's goal. This paper aims to develop and experimentally evaluate an LLM-assisted method that constructs cyber incident scenarios with verified object-event connections and prioritizes them for specialist review in electricity distribution organizations. Achieving this goal involves the following tasks:

1. To develop separate but explicitly connected representations of organizational objects and of the events that lead to the target incident.
2. To implement the generation, validation, refinement and path construction procedures as a software workflow.
3. To evaluate the method on a reference case of a regional electricity distribution operator by comparing generation conditions, ranking the complete scenario variants and identifying the relationships they share.

RESEARCH METHODS

Decision-support workflow

Four stages make up the decision-support workflow for the systematic analysis of cyber incident scenarios:

1. task preparation: the analyst specifies the organization, target incident, assigned dimension and output schema when submitting the task to the Codex desktop application;
2. generation and validation: the structures generated in the OpenAI Codex desktop application pass schema and graph validation;
3. matching and refinement: events are connected to objects, and if the analyst identifies missing details, they return to the generation stage as local refinement tasks;
4. scenario construction and ranking: complete paths are ranked using AHP and examined for shared dependencies as a basis for security review.

Python code implements the software part of the workflow with the jsonschema library for schema validation, NetworkX for graph checks and path search, and NumPy for AHP calculations. A reference model, which remains unavailable to the LLM, is used at the evaluation stage to match generated elements and calculate the refinement priority. The architecture of the proposed workflow is shown in Fig. 1.

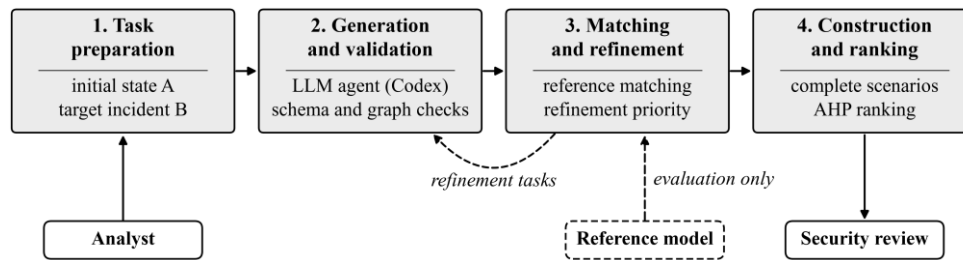


Figure 1. Architecture of the decision-support workflow

Object and event representation

The object hierarchy is rooted at A (the organization in the initial state), while events are generated backward from the target incident B . The combined model of the two views is defined as:

$$M = (V_O, V_E, E_S, E_C, E_M) \quad (1)$$

where V_O and V_E are the sets of persistent objects and state-changing events;
 E_S and E_C contain the structural and directed causal relationships, respectively;
 E_M connects events to objects as actors, targets, channels or resources.

Both structures complement each other, as structural relationships describe composition, interfaces and dependencies, while backward event generation identifies possible causes of the target incident. Although event generation proceeds from effect to cause, the stored causal edges run forward in time. Graph checks also test the graphs for cycles and the object hierarchy for connectivity.

Scenario verification checks whether each event has organizational support by classifying a matched object-event connection as verified only when its object is connected to the organizational root. A scenario is complete when every event has a verified connection and the events form a continuous chain ending at B . JSON validation cannot confirm such properties, which require graph and connection checks.

Generation prompt and output schema

Object and event generation uses a common instruction (Fig. 2) whose role and dimension placeholders adapt the task to the six dimensions of the organization.

The output schema stores objects and events in separate lists to keep the organizational structure distinct from the event sequence. Parent-child and cause-effect fields define the structural and causal relationships. A list of required roles specifies the supporting objects of each event. Refinement tasks also include the gap identified by the analyst and the related elements. The main fields of the schema are listed below:

```

"required": ["case_id", "run_id", "role_id", "role_label",
  "primary_dimension", "object_nodes", "object_edges", "event_nodes",
  "causal_edges"],
"object_edge": {"required": ["parent", "child", "relation", "rationale"]},
"event_node": {"required": ["id", "label", "node_type", "event_class",
  "domain", "required_roles", "predecessor_logic", "observable_result"]},
"causal edge": {"required": ["cause", "effect", "relation", "rationale"]}
  
```

You are the {{ROLE_LABEL}} agent. Read scenario/scenario_brief.json and concentrate on {{PRIMARY_DIMENSION}}.

Construct two trees independently.

Left: initial-state objects

Start at object root A and decompose what exists in the initial state. Every node is a noun phrase: a domain, role, identity, service, system, interface, dependency, channel, power resource, or response resource. Store availability, trust, and connectivity as attributes, not events. Use only structural parent-child edges. Produce at least three levels below A and one real branch.

Right: preceding causal events

Start at event root B and repeatedly ask which earlier event could have caused, enabled, or contributed to it. Every label is a finite verb clause. Produce at least three causal levels and one genuine alternative branch. Generation moves from effect to causes, but store edges with explicit cause and effect fields in chronological orientation.

Do not write a complete A-to-B scenario, link the trees, or treat a structural edge as causal. The connector will later test object-event compatibility.

Return one JSON object conforming to schema/role_tree.schema.json. Use A only as the left root and B only as the right root. Keep content organizational and non-operational; omit products, addresses, credentials, commands, payloads, and exploitation instructions.

Figure 2. Stored prompt template for separate object and event generation



Scenario ranking

Four criteria determine which of the constructed scenarios should be prioritized for review: likelihood, path directness, attack feasibility and detection complexity. Their weights are obtained as the principal eigenvector of a matrix of pairwise comparisons, which express the relative importance of the criteria [15].

The analyst rates likelihood, attack feasibility and detection complexity on a common 1–9 scale. Path directness is rated by the path length:

- likelihood (L): ratings of 1–3, 4–6 and 7–9 correspond to weak evidence, evidence that depends on additional assumptions, and strong confirmation by the case or sources;
- path directness (D): 9 for five or fewer steps between events on the path to the incident, 7 for six or seven, 5 for eight and 3 for nine or more;
- attack feasibility (F): the ratings mean high effort or difficult preconditions, moderate effort with access requirements, and easy entry through exposed or trusted access paths;
- detection complexity (C): the ratings indicate strong early signals, partial visibility and limited early visibility.

Before weighting, the internal consistency of the comparisons is checked using:

$$CI = \frac{\lambda_{\max} - n}{n - 1}, \quad CR = \frac{CI}{RI_n} \quad (2)$$

where $\lambda_{\max}$ is the largest eigenvalue, n is the number of criteria and RI_n is the corresponding random index. For four criteria, $RI_4 = 0.90$.

Dividing each rating by 9, multiplying it by the criterion weight and summing the contributions gives the priority score:

$$S_i = \sum_{j=1}^n \alpha_j \frac{r_{ij}}{9} \quad (3)$$

where S_i is the score of variant i , r_{ij} is its rating for criterion j , and α_j is the criterion weight. This score cannot exceed 1 because the four weights sum to 1. For comparison, the minimum path length criterion selects the variant with the fewest steps between events.

Computer experiment

Reference case. The complete workflow is tested on a reference case that models a Ukrainian regional electricity distribution organization with generalized roles and assets. This organization is based on the reports of the 2015 and 2022 incidents [2]–[4]. Following NIST SP 800-82 Rev. 3 [20], it distinguishes enterprise IT, the boundary with operational technology (OT), operational control and physical processes.

In the initial state, operation is normal: enterprise and OT services are available, substations are energized, approved resources exist for administration, maintenance and response, and dispatchers can monitor and control the system. Under such conditions, the target incident is specified as unauthorized switching that de-energizes multiple substations.

At the preparation stage, the analyst decomposes the organization into six dimensions. Their refinement weights are as follows: identity and personnel (0.16), enterprise IT and remote administration (0.18), IT-OT interfaces and access (0.18), operational control and physical process (0.25), continuity and response resources (0.13), as well as external and supplier dependencies (0.10). Operational control has the highest weight because it is closest to switching. Enterprise IT and IT-OT access follow because of the attack routes documented in the incident reports.

Experimental setup. Object and event structures are generated by the GPT-5.6 Sol model (identifier gpt-5.6-sol) at the ultra reasoning effort setting in the OpenAI Codex desktop application and analyzed in Python. Overall, the asymmetric procedure includes 11 sessions with 18 outputs: six in three initial sessions and 12 in eight refinement sessions over four rounds. The direct baseline, in which objects and events are generated jointly, consists of one session with one output under the same organization, dimensions, model setting and validation rules. Since both conditions use a single model setting, the differences between them result from the generation procedure.

Validation and reference access. Session inputs are limited to a description of the organization, the prompt, the schema, permitted earlier outputs and identified refinement gaps. According to the session records, the reference model and the node mapping were not accessed at this stage. Thus, the reference model affected generation only through the selected refinement gaps. For acceptance, all outputs are checked against the JSON schema (Draft 2020-12) and graph rules for node types, endpoints, permitted edge types and acyclicity.

Reference matching. Recovery is assessed by comparing the generated elements with the fixed reference model, which contains 40 objects, 39 structural relationships, 16 events, 16 causal relationships and 16 object-event connections. Generated nodes are matched to reference nodes by type, domain and defining attributes in



separate Codex sessions that accept only equivalent nodes as matches. Broader, narrower or combined nodes remain unmatched. This recorded mapping was applied automatically. For the direct output, the author verified the mapping of all 41 generated nodes. An edge is counted as recovered when its mapped source and target match those of a reference edge. Precision is measured as the share of generated nodes with a reference equivalent. The event sequences of the five reference variants, fixed in advance of the experiment, define which variants count as complete throughout refinement.

For local refinement, recovery of the object structure, event structure and verified connections is measured against the reference model within each dimension, and the refinement priority is calculated as:

$$G_d = w_d [1 - \min(C_d^O, C_d^E, C_d^G)] \quad (4)$$

Here, d denotes a dimension and w_d is its refinement weight. C_d^O is the recovered share of object nodes and structural relationships, C_d^E that of event nodes and causal relationships, and C_d^G that of verified object-event connections. A larger G_d indicates a more urgent need for local expansion. In practice, it is necessary to replace the reference model in this calculation with domain checklists and a review of unresolved connections.

Judgments and evaluation. The ratings and pairwise criterion comparisons were made by external domain analysts who had reviewed the organization, the complete paths and the incident reports. Table 1 contains the resulting comparison matrix and weights.

Table 1

Pairwise criterion comparisons and AHP weights

Criterion	L	D	F	C	Weight
Likelihood (L)	1	3	3	2	0.459
Path directness (D)	1/3	1	1	1/2	0.144
Attack feasibility (F)	1/3	1	1	1	0.172
Detection complexity (C)	1/2	2	1	1	0.226

In the matrix, the value 3 for likelihood relative to directness has the inverse value 1/3. A value of 1 means equal importance. Equation (2) gives a consistency index of 0.015 for the maximum eigenvalue of 4.046 and a consistency ratio of 0.017, which is below the conventional threshold of 0.10. Scores are calculated with exact, not rounded, weights.

To test how sensitive the AHP ranking of the five variants is to the criterion weights, 10,000 runs were performed with fixed paths and ratings. In these runs, each weight is independently multiplied by a uniformly distributed factor in the range 0.8–1.2, and the weights are then normalized to sum to 1.

THE RESULTS AND DISCUSSION

Scenario recovery and local refinement

Figure 3 shows the recovered shares of the object structure, event structure and verified connections after direct and initial asymmetric scenario synthesis and after the four refinement rounds. Table 2 lists the corresponding numbers of elements. Initially, the asymmetric generation recovers the same six events as the direct baseline but more objects and structural relationships (22 and 15 compared with 17 and 13). For example, the direct output contains the event “Supplier-mediated access traverses the controlled IT-OT interface”, which leads directly to operational control without the administrative access to the OT support environment and to the dispatch services required by the reference model. Complete scenarios are built in the refinement rounds, because the event chains of both initial outputs are not yet connected to supporting objects.

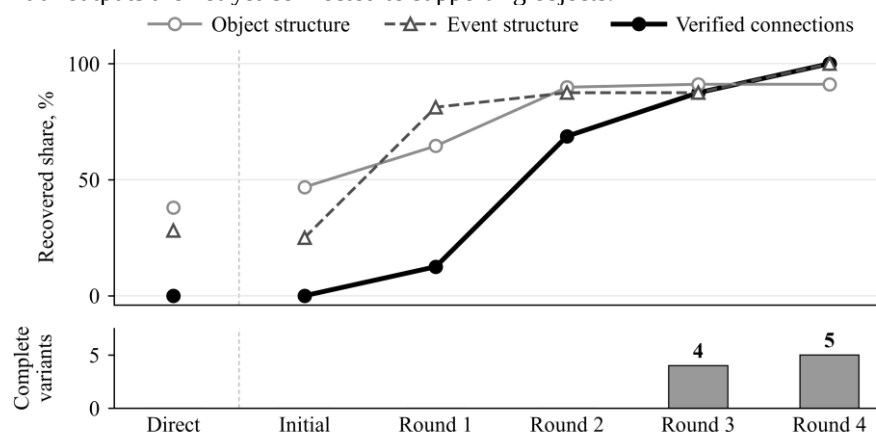


Figure 3. Recovery of the reference model across generation conditions

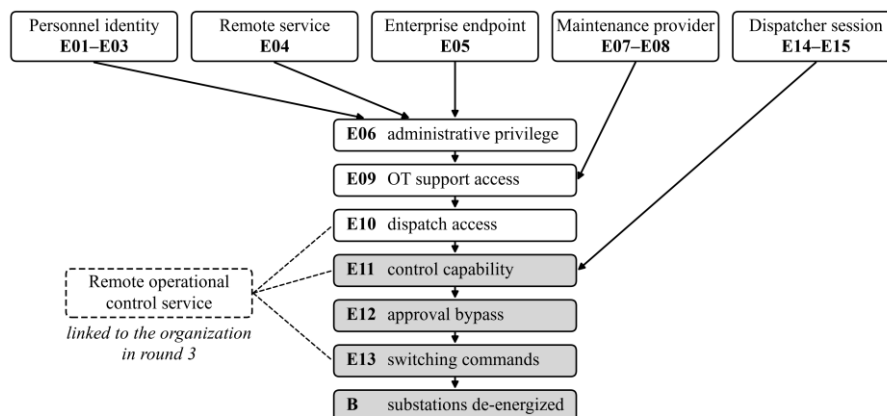


Table 2

Recovered reference elements after direct and initial asymmetric generation, and after refinement rounds 1–4

Reference element	Direct	Initial	Round 1	Round 2	Round 3	Round 4
Object nodes (of 40)	17	22	30	37	37	37
Structural relationships (of 39)	13	15	21	34	35	35
Event nodes (of 16)	6	6	14	14	14	16
Causal relationships (of 16)	3	2	12	14	14	16
Verified object-event connections (of 16)	0	0	2	11	14	16
Complete scenario variants (of 5)	0	0	0	0	4	5

According to Fig. 3, step-by-step refinement of organizational relationships and causal events completes the scenarios. It expands both structures to 37 objects and 14 events during the first two rounds. Most of the growth occurs in round 2, when structural relationships increase from 21 to 34 and verified object-event connections from 2 to 11. The remote operational control service, however, remains disconnected from the organizational hierarchy until its missing relationship is restored in round 3, which completes four variants. The fifth variant (personnel credential compromise) is completed in round 4 through two personnel entry events that were added to the asymmetric structures. At that point, all events, causal relationships and object-event connections of the reference model are recovered. These rounds also show that four variants depend on a single organizational relationship rather than on additional events. Figure 4 shows how the attack paths converge and how they depend on the shared operational control service.

*Figure 4. Convergence of the attack paths of the five scenario variants*

The attack paths from all five entry routes converge on one chain that ends in the target incident B and contains the shaded events common to the five variants. Four of the variants also pass through E09→E10→E11 and meet at the operational control service on the way to B. Dashed lines show that events E10, E11 and E13 depend on this service. In addition, the reference model contains a direct link E11→E13 for commands issued without the operator workstation. None of the five variants uses this link.

Precision and cost are also compared for the two conditions. Excluding the given endpoints A and B, direct generation produces 21 objects and 20 events. Sixteen of the objects and nine of the events have a reference equivalent. These nine events correspond to five reference events, so the 17 objects and six events in Table 2 include the two endpoints. After round 4, the asymmetric structures have equivalents for 76 of 181 objects and 50 of 156 events. Among the unmatched elements, 77 objects are parts of matched objects and 49 events are directly linked to matched events. Forty-five of the remaining 57 events describe security measures, response actions or supplier controls that fail to stop the attack (for example, “Boundary monitoring fails to detect the anomalous session”). Such events lie outside the reference model, because it contains only attack steps. Asymmetric sessions lasted 3.0–13.8 minutes each, 89.5 minutes in total, including one review of an intermediate draft in round 2. Altogether, 8.5 million tokens were processed, about 10 times as many as in the direct session, which took 14.6 minutes and used 0.83 million tokens.



Completed scenarios and ranking

Table 3 associates each event of the personnel credential compromise scenario with the required object, listing the relevant security domain or MITRE ATT&CK technique for Enterprise or for industrial control systems (ICS) [21], [22]. Its event identifiers follow Fig. 4, where events E04–E05, E07–E08 and E14–E15 belong to other variants.

This scenario leads from a targeted message and compromised administrator credentials (E01–E03) through enterprise privilege and IT-OT access (E06, E09, E10) to the control service and operator workstation used for unauthorized switching (E11–E13). Every step of the scenario has a supporting object. Four events depend on enterprise administrator objects and three on the remote operational control service, which links the IT-OT access stage to the switching stages.

Table 3

Events and supporting objects of the personnel credential compromise scenario

ID	Event	Required object	Domain / ATT&CK
E01	A targeted message reaches a user with enterprise administration rights	Enterprise administrator role (target)	Identity/personnel; T1566 Enterprise
E02	Enterprise administrator credentials or a valid session token is leaked	Enterprise administrator account and authentication material	Identity/personnel
E03	An unauthorized valid account session is established	Enterprise administrator accounts (target)	Identity/personnel; T1078 Enterprise
E06	Enterprise administrative privilege is obtained without authorization	Enterprise administrator accounts (resource)	Enterprise IT; T1078 Enterprise
E09	Administrative access to the OT support environment is obtained	OT support environment (target)	IT-OT access; T0859 ICS
E10	Unauthorized access reaches dispatch and control services	Remote operational control service (target)	IT-OT access; T0886 ICS
E11	Unauthorized operational control capability is obtained	Remote operational control service (resource)	Operational control
E12	An operator workstation accepts a switching action without a dispatcher order or switching form	Operator workstation	Operational control; T0823 ICS
E13	Unauthorized switching commands are issued	Remote operational control service (resource)	Operational control; T1692.001 ICS
B	Unauthorized switching de-energizes multiple substations	Distribution substations (target)	Physical process; T0831, T0826 ICS

Table 4 gives the scenario lengths, criterion ratings and resulting AHP scores. Its Steps column shows the path length in steps between events. The criterion columns contain the ratings for likelihood (L), path directness (D), attack feasibility (F) and detection complexity (C).

Table 4

Scenario lengths, criterion ratings and AHP scores

Scenario variant	Steps	L	D	F	C	Score
Personnel credential compromise	9	9	3	7	7	0.816
Dispatcher session takeover	5	7	9	5	6	0.746
Remote service compromise	7	7	7	7	5	0.728
Maintenance provider credential compromise	7	5	7	7	8	0.701
Enterprise endpoint compromise	7	7	7	6	4	0.683

According to Table 4, personnel credential compromise has the highest likelihood, because the incident reports document targeted contact, compromised administrative identities and activity in valid sessions [2], [3].



Dispatcher session takeover is shorter but is rated as more demanding, because the adversary must abuse a dispatcher account authorized for remote control and take over an operator session. Incident reports give weaker evidence for the provider route. Trusted maintenance access, however, makes this route feasible and its early detection difficult. In 2024, CERT-UA reported that supplier access was used to prepare sabotage at about 20 energy, water and heat supply facilities [23]. A likelihood rating of 7 would move the provider route to second place with a score of 0.803, leaving personnel credential compromise first in 80.1% of the sensitivity runs.

Remote service compromise receives ratings of 7 for likelihood and feasibility because the exposed administration service provides easy entry leading to the OT environment. Boundary monitoring lowers the detection complexity of this variant to 5. Enterprise endpoint compromise receives a likelihood rating of 7 for documented entry, a feasibility rating of 6 for moderate effort and a detection complexity rating of 4, which reflects endpoint and authentication monitoring. All three variants with seven steps receive a directness rating of 7.

These results show that the multi-criteria ranking differs from a ranking by the minimum path length alone. Personnel credential compromise receives the highest AHP score (0.816) and keeps first place in all sensitivity runs with varied weights. Ranking by minimum path length selects dispatcher session takeover instead. Path length alone is therefore insufficient for prioritization, since the AHP ranking reflects the importance of each scenario under the selected criteria. The first two positions do not change in the sensitivity runs, while the positions of the lower three variants change in 3.2–19.5% of them. No one-point change in one rating alters the leading variant, but a likelihood rating of 7 instead of 9 would lower its score to 0.714, below that of dispatcher session takeover (0.746).

Attack trees in [8]–[10] are constructed for a predefined system model or specification. The proposed synthesis builds the organizational structure and the event chains itself and verifies how they are connected. It differs from [16] in ranking the alternatives it generates.

Use in security review

Personnel credential compromise, the top-ranked variant, makes administrator credentials, exposure to targeted messages and privileged sessions the main priorities for security review. Dispatcher session takeover draws attention to operator accounts and active sessions. The other variants point to remote administration, supplier maintenance and enterprise endpoints.

Critical dependencies shared by several scenarios were identified by removing individual relationships. In the reference model, all five paths share three causal relationships, which link the control capability, the approval bypass, the unauthorized commands and the target incident in sequence. No accepted path remains if even one of these relationships is removed. Removing the two relationships leading to OT support access (E06→E09 and E08→E09) eliminates four variants, leaving only dispatcher session takeover. If the object-event connection of any single entry event is removed, the variant of that event alone is eliminated. It is therefore necessary to verify the account, service or channel of each entry route separately.

The synthesized scenarios are also relevant to national cybersecurity rules, since Order No. 403 of the Ministry of Energy of Ukraine requires threats and their potential consequences to be documented (ID.RA-03, ID.RA-04) [24]. The object-event connections that link the synthesized events to persistent organizational objects can be used to document the affected assets and possible consequences. Table 3 can also be used in joint exercises for security operations center (SOC) analysts and dispatch personnel. In such exercises, participants can check switching authority and operational records under the Rules for Performing Operational Switching in Electrical Installations [25].

CONCLUSIONS AND PROSPECTS FOR FURTHER RESEARCH

This study developed and experimentally evaluated a method for LLM-assisted synthesis and AHP prioritization of cyber incident scenarios. The scientific novelty of the study is the asymmetric synthesis of scenarios. This synthesis decomposes the organization in the initial state into organizational objects and generates causal event chains backward from the target incident. An event sequence is accepted as a scenario only when every supporting object is connected to the organizational root. In [17]–[19], LLMs estimated weights, proposed criteria and alternatives, or judged security controls. Unlike these studies, the synthesis constructs verified scenario alternatives together with their supporting organizational objects as input for multi-criteria prioritization.

In the experiment, four refinement rounds recovered all 16 events, 16 causal relationships and 16 object-event connections of the reference model, completing five variants. These results confirm that four of them depend on a single relationship between the shared control service and the organizational hierarchy and that the fifth requires further entry events. Personnel credential compromise was first in the AHP ranking with a score of 0.816 and retained this place in all 10,000 sensitivity runs, whereas the minimum path length criterion selected dispatcher session takeover.

The results are of practical significance, since they identify the entry routes and the three causal



relationships shared by all scenario paths. These findings suggest that electricity distribution operators should give priority to protecting administrator credentials, privileged sessions and remote IT-OT and supplier access. Switching orders should be verified at operator workstations, and switching commands should be monitored. Such results may be useful for documenting threats and consequences under Order No. 403.

The method was evaluated on one reference case with one generation sequence per condition. Further research will extend the evaluation to other critical infrastructure cases and other LLMs with repeated runs and will compare generation conditions under equal effort. Checklist-based refinement will be tested with operational specialists.

Declaration on the use of artificial intelligence. GPT-5.6 Sol in the OpenAI Codex desktop application was the object of the experiment. It generated the analyzed JSON structures, mapped them to the reference model in separate sessions and assisted with code drafting. Claude (Opus 5.5) assisted with the figures and the precision and cost analyses from the stored results. Gemini (3.1 Pro) was used for language editing and review. The author designed the method and the experiment and defined the case, reference model and acceptance rules. The author also checked the node mapping of the direct output and verified all outputs against the stored experimental records. Full responsibility for the content rests with the author.

Conflict of interest. The author declares no conflicts of interest.

Funding. This research received no external funding.

Data availability. The case description, reference model, prompts and scripts are available from the author upon request. The generated attack scenarios are not published openly because they describe attack paths against critical infrastructure.

REFERENCES

1. State Service of Special Communications and Information Protection of Ukraine. (2026, January 7). *CERT-UA processed nearly 6,000 cyber incidents in 2025: Hostile attack volume rose by 37%*. <https://cip.gov.ua/en/news/cert-ua-u-2025-roci-opracyuvala-maizhe-6000-kiberincidentiv-kilkist-vorozhikh-atak-zrosla-na-37>
2. Lee, R. M., Assante, M. J., & Conway, T. (2016, March 18). *Analysis of the cyber attack on the Ukrainian power grid: Defense use case*. Electricity Information Sharing and Analysis Center. <https://nsarchive.gwu.edu/sites/default/files/documents/3891751/SANS-and-Electricity-Information-Sharing-and.pdf>
3. Maynard, P., McLaughlin, K., & Sezer, S. (2020). Decomposition and sequential-AND analysis of known cyber-attacks on critical infrastructure control systems. *Journal of Cybersecurity*, 6(1), Article tyaa020. <https://doi.org/10.1093/cybsec/tyaa020>
4. State Service of Special Communications and Information Protection of Ukraine. (2022, April 12). *Heavy cyberattack on Ukraine's energy sector prevented*. <https://cip.gov.ua/en/news/poperedzhena-masshtabna-kiberataka-na-energetichnii-sektor-ukrayini>
5. Ferdous, R., Khan, F., Sadiq, R., Amyotte, P., & Veitch, B. (2013). Analyzing system safety and risks under uncertainty using a bow-tie diagram: An innovative approach. *Process Safety and Environmental Protection*, 91(1–2), 1–18. <https://doi.org/10.1016/j.psep.2011.08.010>
6. Jhawar, R., Kordy, B., Mauw, S., Radomirović, S., & Trujillo-Rasua, R. (2015). Attack trees with sequential conjunction. In *ICT Systems Security and Privacy Protection* (pp. 339–353). Springer. https://doi.org/10.1007/978-3-319-18467-8_23
7. International Organization for Standardization. (2024). *Automation systems and integration – Object-Process Methodology* (ISO Standard No. 19450:2024). <https://www.iso.org/standard/84612.html>
8. Bryans, J., Liew, L. S., Nguyen, H. N., Sabaliauskaite, G., Shaikh, S., & Zhou, F. (2020). A template-based method for the generation of attack trees. In *Information Security Theory and Practice* (pp. 155–165). Springer. https://doi.org/10.1007/978-3-030-41702-4_10
9. Groner, R., Witte, T., Raschke, A., Hirn, S., Pekaric, I., Frick, M., Tichy, M., & Felderer, M. (2023). Model-based generation of attack-fault trees. In *Computer Safety, Reliability, and Security* (pp. 107–120). Springer. https://doi.org/10.1007/978-3-031-40923-3_9
10. Birchler De Allende, A., Sultan, B., & Apvrille, L. (2025). Automated attack tree generation using artificial intelligence and natural language processing. In *Risks and Security of Internet and Systems* (pp. 141–156). Springer. https://doi.org/10.1007/978-3-031-89350-6_9
11. Prapty, R. T., Kundu, A., & Iyengar, A. (2024). *Using retriever augmented large language models for attack graph generation* (arXiv:2408.05855). arXiv. <https://doi.org/10.48550/arXiv.2408.05855>
12. Ahmed, C. M. (2025). AttackLLM: LLM-based attack pattern generation for an industrial control system. In *Proceedings of the 2nd International Workshop on Foundation Models for Cyber-Physical Systems &*



- Internet of Things* (pp. 31–36). ACM. <https://doi.org/10.1145/3722565.3727196>
13. Guan, L., Valmeekam, K., Sreedharan, S., & Kambhampati, S. (2023). Leveraging pre-trained large language models to construct and utilize world models for model-based task planning. *Advances in Neural Information Processing Systems*, 36, 79081–79094. <https://doi.org/10.52202/075280-3459>
 14. Andreiev, D., Chornyi, A., Stopochkina, I., & Ilin, M. (2026). Methodology for automating cyber incident reports using LLM. *Cybersecurity: Education, Science, Technique*, 1(33), 274–285. <https://doi.org/10.28925/2663-4023.2026.33.1156>
 15. Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, 48(1), 9–26. [https://doi.org/10.1016/0377-2217\(90\)90057-I](https://doi.org/10.1016/0377-2217(90)90057-I)
 16. Dolgova, N., Shapovalova, O., & Solodovnyk, H. (2025). Decision support framework for critical infrastructure cybersecurity planning. *Cybersecurity: Education, Science, Technique*, 1(29), 348–372. <https://doi.org/10.28925/2663-4023.2025.29.884>
 17. Lande, D., Alekseichuk, L., Svoboda, I., & Strashnoy, L. (2024). Methodology of a swarm of virtual experts for evaluating the weight of connections in networks. *Theoretical and Applied Cybersecurity*, 6(2), 25–33. <https://doi.org/10.20535/tacs.2664-29132024.2.319946>
 18. Svoboda, I. (2025). Automating cybersecurity decision-making with AI and the analytic hierarchy process. *Theoretical and Applied Cybersecurity*, 7(2), 91–97. <https://doi.org/10.20535/tacs.2664-29132025.2.338865>
 19. Svoboda, I., & Lande, D. (2026). Prioritization of security controls for critical infrastructure using the analytic network process and large language models. *Collection "Information Technology and Security"*, 14(1), 235–252. <https://doi.org/10.20535/2411-1031.2026.14.1.365500>
 20. Stouffer, K., Pease, M., Tang, C., Zimmerman, T., Pillitteri, V., Lightman, S., Hahn, A., Saravia, S., Sherule, A., & Thompson, M. (2023). *Guide to Operational Technology (OT) Security* (NIST Special Publication 800-82 Revision 3). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-82r3>
 21. MITRE ATT&CK. (2026). *Enterprise Matrix* (ATT&CK version 19.2). <https://attack.mitre.org/versions/v19/matrices/enterprise/>
 22. MITRE ATT&CK. (2026). *ICS Matrix* (ATT&CK version 19.2). <https://attack.mitre.org/versions/v19/matrices/ics/>
 23. Government Computer Emergency Response Team of Ukraine. (2024, April 19). *Plans of UAC-0133 (Sandworm) for cyber sabotage at almost 20 critical infrastructure facilities of Ukraine*. <https://cert.gov.ua/article/6278706>
 24. Ministry of Energy of Ukraine. (2025). *On approval of the Cyber protection requirements for the fuel and energy sector of critical infrastructure* (Order No. 403). <https://zakon.rada.gov.ua/go/z1599-25>
 25. Ministry of Energy and Coal Industry of Ukraine. (2018). *On approval of the Rules for performing operational switching in electrical installations* (Order No. 77). <https://zakon.rada.gov.ua/go/z0211-18>

**Свобода Ігор Максимович**

аспірант кафедри інформаційної безпеки

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна

ORCID: 0009-0001-1709-2904

i.svoboda@kpi.ua

МЕТОД СИНТЕЗУ ТА ПРІОРИТИЗАЦІЇ СЦЕНАРІЇВ КІБЕРІНЦИДЕНТІВ У КРИТИЧНІЙ ІНФРАСТРУКТУРІ З ВИКОРИСТАННЯМ LLM І МАІ

Анотація. У статті розглянуто, як пов'язати кожен етап багатетапної кібератаки на об'єкти критичної інфраструктури з активами, обліковими записами та системами оперативного керування, залученими до нього. Наявні методи на основі дерев атак потребують наперед заданої моделі системи, а підходи на основі великих мовних моделей (LLM) не перевіряють, чи спирається кожен згенерований етап атаки на організаційні об'єкти. Дослідження має на меті розробити й експериментально перевірити метод підтримки прийняття рішень на основі LLM для автоматизованого синтезу та багатокритеріальної пріоритизації сценаріїв кіберінцидентів. Запропонований метод ґрунтується на асиметричній генерації: організацію в початковому стані функціонування ієрархічно декомпонують на активи, а причинно-наслідкові послідовності подій будують у зворотному напрямку від заданого інциденту. Згенеровані LLM структури перевіряють за JSON-схемою та на узгодженість графів, а аналітик визначає відсутні елементи для локального уточнення. Завершені сценарії ранжують за пріоритетом методом аналізу ієрархій (МАІ). Послідовність подій вважають сценарієм лише тоді, коли кожен її опорний об'єкт належить до організаційної ієрархії. Для перевірки методу використано змодельований випадок несанкціонованого перемикання, яке знеструмує підстанції регіонального оператора системи розподілу електроенергії в Україні. За чотири раунди уточнення відновлено всі події еталонної моделі разом із їхніми причинно-наслідковими зв'язками та зв'язками з об'єктами, а також побудовано п'ять завершених варіантів сценаріїв атак. Встановлено, що чотири з них залежали від одного зв'язку між спільним сервісом керування та організаційною ієрархією. За результатами МАІ сценарій компрометації облікових даних персоналу отримав найвищий пріоритет з оцінкою 0,816 і випередив коротший сценарій захоплення сеансу диспетчера. Перше місце він зберіг в усіх прогонах аналізу чутливості. Отримані результати підтверджують, що поєднання LLM із формальними перевітками графів виявляє етапи атаки, не підкріплені організаційними об'єктами. На практиці центри операцій безпеки (SOC) можуть використовувати підхід, щоб визначати критичні залежності атак і вирішувати, які облікові дані та оперативні процедури захищати насамперед.

Ключові слова: великі мовні моделі; сценарії кіберінцидентів; синтез об'єктів і подій; метод аналізу ієрархій; критична інфраструктура; розподіл електроенергії; підтримка прийняття рішень.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. State Service of Special Communications and Information Protection of Ukraine. (2026, January 7). *CERT-UA processed nearly 6,000 cyber incidents in 2025: Hostile attack volume rose by 37%*. <https://cip.gov.ua/en/news/cert-ua-u-2025-roci-opracuyvala-maizhe-6000-kiberincidentiv-kilkist-vorozhikh-atak-zroslo-na-37>
2. Lee, R. M., Assante, M. J., & Conway, T. (2016, March 18). *Analysis of the cyber attack on the Ukrainian power grid: Defense use case*. Electricity Information Sharing and Analysis Center. <https://nsarchive.gwu.edu/sites/default/files/documents/3891751/SANS-and-Electricity-Information-Sharing-and.pdf>
3. Maynard, P., McLaughlin, K., & Sezer, S. (2020). Decomposition and sequential-AND analysis of known cyber-attacks on critical infrastructure control systems. *Journal of Cybersecurity*, 6(1), Article tyaa020. <https://doi.org/10.1093/cybsec/tyaa020>
4. State Service of Special Communications and Information Protection of Ukraine. (2022, April 12). *Heavy cyberattack on Ukraine's energy sector prevented*. <https://cip.gov.ua/en/news/poperedzhena-masshtabna-kiberataka-na-energetichnii-sektor-ukrayini>
5. Ferdous, R., Khan, F., Sadiq, R., Amyotte, P., & Veitch, B. (2013). Analyzing system safety and risks under uncertainty using a bow-tie diagram: An innovative approach. *Process Safety and Environmental Protection*, 91(1–2), 1–18. <https://doi.org/10.1016/j.psep.2011.08.010>



6. Jhawar, R., Kordy, B., Mauw, S., Radomirović, S., & Trujillo-Rasua, R. (2015). Attack trees with sequential conjunction. In *ICT Systems Security and Privacy Protection* (pp. 339–353). Springer. https://doi.org/10.1007/978-3-319-18467-8_23
7. International Organization for Standardization. (2024). *Automation systems and integration – Object-Process Methodology* (ISO Standard No. 19450:2024). <https://www.iso.org/standard/84612.html>
8. Bryans, J., Liew, L. S., Nguyen, H. N., Sabaliauskaite, G., Shaikh, S., & Zhou, F. (2020). A template-based method for the generation of attack trees. In *Information Security Theory and Practice* (pp. 155–165). Springer. https://doi.org/10.1007/978-3-030-41702-4_10
9. Groner, R., Witte, T., Raschke, A., Hirn, S., Pekaric, I., Frick, M., Tichy, M., & Felderer, M. (2023). Model-based generation of attack-fault trees. In *Computer Safety, Reliability, and Security* (pp. 107–120). Springer. https://doi.org/10.1007/978-3-031-40923-3_9
10. Birchler De Allende, A., Sultan, B., & Aprville, L. (2025). Automated attack tree generation using artificial intelligence and natural language processing. In *Risks and Security of Internet and Systems* (pp. 141–156). Springer. https://doi.org/10.1007/978-3-031-89350-6_9
11. Prapty, R. T., Kundu, A., & Iyengar, A. (2024). *Using retriever augmented large language models for attack graph generation* (arXiv:2408.05855). arXiv. <https://doi.org/10.48550/arXiv.2408.05855>
12. Ahmed, C. M. (2025). AttackLLM: LLM-based attack pattern generation for an industrial control system. In *Proceedings of the 2nd International Workshop on Foundation Models for Cyber-Physical Systems & Internet of Things* (pp. 31–36). ACM. <https://doi.org/10.1145/3722565.3727196>
13. Guan, L., Valmeekam, K., Sreedharan, S., & Kambhampati, S. (2023). Leveraging pre-trained large language models to construct and utilize world models for model-based task planning. *Advances in Neural Information Processing Systems*, 36, 79081–79094. <https://doi.org/10.52202/075280-3459>
14. Andreiev, D., Chorneyi, A., Stopochkina, I., & Ilin, M. (2026). Methodology for automating cyber incident reports using LLM. *Cybersecurity: Education, Science, Technique*, 1(33), 274–285. <https://doi.org/10.28925/2663-4023.2026.33.1156>
15. Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, 48(1), 9–26. [https://doi.org/10.1016/0377-2217\(90\)90057-1](https://doi.org/10.1016/0377-2217(90)90057-1)
16. Dolgova, N., Shapovalova, O., & Solodovnyk, H. (2025). Decision support framework for critical infrastructure cybersecurity planning. *Cybersecurity: Education, Science, Technique*, 1(29), 348–372. <https://doi.org/10.28925/2663-4023.2025.29.884>
17. Lande, D., Alekseichuk, L., Svoboda, I., & Strashnoy, L. (2024). Methodology of a swarm of virtual experts for evaluating the weight of connections in networks. *Theoretical and Applied Cybersecurity*, 6(2), 25–33. <https://doi.org/10.20535/tacs.2664-29132024.2.319946>
18. Svoboda, I. (2025). Automating cybersecurity decision-making with AI and the analytic hierarchy process. *Theoretical and Applied Cybersecurity*, 7(2), 91–97. <https://doi.org/10.20535/tacs.2664-29132025.2.338865>
19. Svoboda, I., & Lande, D. (2026). Prioritization of security controls for critical infrastructure using the analytic network process and large language models. *Collection "Information Technology and Security"*, 14(1), 235–252. <https://doi.org/10.20535/2411-1031.2026.14.1.365500>
20. Stouffer, K., Pease, M., Tang, C., Zimmerman, T., Pillitteri, V., Lightman, S., Hahn, A., Saravia, S., Sherule, A., & Thompson, M. (2023). *Guide to Operational Technology (OT) Security* (NIST Special Publication 800-82 Revision 3). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-82r3>
21. MITRE ATT&CK. (2026). *Enterprise Matrix (ATT&CK version 19.2)*. <https://attack.mitre.org/versions/v19/matrices/enterprise/>
22. MITRE ATT&CK. (2026). *ICS Matrix (ATT&CK version 19.2)*. <https://attack.mitre.org/versions/v19/matrices/ics/>
23. Government Computer Emergency Response Team of Ukraine. (2024, April 19). *Plans of UAC-0133 (Sandworm) for cyber sabotage at almost 20 critical infrastructure facilities of Ukraine*. <https://cert.gov.ua/article/6278706>
24. Ministry of Energy of Ukraine. (2025). *On approval of the Cyber protection requirements for the fuel and energy sector of critical infrastructure* (Order No. 403). <https://zakon.rada.gov.ua/go/z1599-25>
25. Ministry of Energy and Coal Industry of Ukraine. (2018). *On approval of the Rules for performing operational switching in electrical installations* (Order No. 77). <https://zakon.rada.gov.ua/go/z0211-18>
1. This work is licensed under Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Отримано редакцією журналу / Received: 16.09.26

Прорецензовано / Revised: 24.09.26

Схвалено до друку / Accepted: 26.10.26



This work is licensed under Creative Commons Attribution-noncommercial-sharealike 4.0 International License.