GenAttackTracker: Real-Time SCADA-based Cyber Threat Detection Through Scoring and Bayesian Model Integration

Fatemeh Movafagh and Uwe Glässer
School of Computing Science, *Simon Fraser University*
British Columbia, Canada
Email: {fma44, glaesser}@sfu.ca

***Abstract*—The increasing sophistication and evolving nature of cyber threats pose significant risks to critical infrastructure systems. This research introduces GenAttackTracker, a novel algorithmic framework designed for real-time detection and interpretation of cyber threats in Supervisory Control and Data Acquisition (SCADA) systems. By integrating dynamic anomaly scoring with hierarchical Bayesian modeling, GenAttackTracker enhances situational awareness for identifying potential security breaches in operational technology environments. This robust mechanism contributes directly to enhancing cyber resilience by improving threat detection in critical infrastructure systems, an essential component of ensuring the continuity and security of mission-critical processes. The framework leverages primary data from SCADA systems and secondary contextual data sources, termed Suspicious Activity Markers (SAMs). Through Bayesian inference, the model continuously updates its understanding of the system's security status, allowing informed decision-making.**

Keywords*—*Cyber Resilience; Critical Infrastructure Security; Cyber-Physical Systems; Supervisory Control and Data Acquisition (SCADA); Online Threat Detection; Bayesian Inference, Anomaly Detection; Suspicious Activity Markers (SAMs); Machine Learning; Real-Time Cyber Threat Detection.

I. Introduction

Cyber threats continually evolve, exerting new capabilities and enhancing their metamorphic nature to evade detection by legacy antivirus products. With evermore sophisticated threats, such as malware-free intrusions and zero-day exploits targeting Critical Infrastructure (CI), cybersecurity breaches become more inevitable—leaving infrastructures on which we all depend at high risk of global threat activity [1]. This reality amplifies fears of catastrophic events tailored to incapacitate CI systems in key infrastructure sectors.

The research project presented here aims at enhancing CI protection by reinforcing security and resilience of mission-critical Operational Technology (OT) against advanced cyber threats. OT is vital to industrial process automation as used for many types of CI facilities, which are often highly interconnected, mutually dependent systems [2]. In manufacturing and production, process automation frequently hinges on OT to interoperate with the physical environment, where Industrial Control Systems (ICS) monitor and control physical processes, devices, and infrastructures. Most prominently, Supervisory Control and Data Acquisition (SCADA) architectures allow large-scale processes to span multiple sites and work over large distances. A SCADA-based OT system is a Cyber-Physical System (CPS) that enables supervisory process control by capturing real-time data of the infrastructure's operational status. Industry sectors using SCADA include manufacturing, oil and natural gas, electrical generation and distribution, maritime, rail, and utilities [3].

With the paradigm shift to Industry 4.0, intelligent process control aims at even tighter integration of digital control loops powered by AI, embedded computing, robotics, and Internet of Things (IoT) with technical processes in the physical environment. This trend inevitably increases fragility of process automation, making OT more vulnerable by amplifying the risk of cascading and escalating failures. Beyond extensive disruptions of critical services, highly orchestrated attacks can result in disastrous physical damage caused by triggering cascading malfunctions to overload mission-critical system components.

Considering that complete security of network technology may be unattainable, the focus shifts to risk mitigation and remediation. Our work aims at proactive measures that reduce the likelihood and the potential impact of severe cyberattacks. Traditional risk mitigation methods are often inadequate for addressing advanced cyber threats due to their highly sophisticated and evolving nature. The gravity of this situation calls for advanced analytical models and algorithmic methods to ensure that cyber situational awareness keeps pace with the evolving threat landscape. Artificial Intelligence (AI) is instrumental in detecting and interpreting abnormal OT system behavior by continuously analyzing supervisory control data streamed from system operations. Abnormal behavior patterns can signal imminent threat activity after a security breach. A timely response launching countermeasures is critical to contain any intrusion before it can spread laterally across wider networks. The dynamics and anatomy of intricate attack scenarios requires advanced analytical models and algorithmic methods for turning cyber situational awareness into actionable intelligence in real-time.

Research Question. For OT systems relying on supervisory control system architectures, such as SCADA, we consider the following research question: *how can contextual data and information from secondary threat intelligence sources substantiate evidence of changes in the system's security status derived from online analysis of control data?* Fusing data and information from a number of causally related events may arguably result in more accurate situational awareness as baseline for online inference and decision-making processes.

Methodology. Inevitable uncertainty due to lack of ground truth is problematic for the reliable detection and interpretation of unexpected behavior patterns relevant a system's security status and also increases the rate of false positives. A Bayesian modeling approach can significantly improve the outcome. Bayesian inference promotes frequent updating of conditional probabilities as new information becomes available, providing a dynamic perspective of security threat levels. Thus, the more technical question is: *how can Bayesian inference and the integration of contextual data and information, termed Suspicious Activity Markers (SAMs), enhance situational awareness of cyber threat activities targeting the operation of CI systems?*

Contribution. The novel contribution of this paper is GenAttackTracker, a generic analytical framework for online detection and interpretation of abnormal behavior patterns in supervisory control data streamed from a mission-critical OT system. Combining dynamic attack scoring with Bayesian inference to fuse results from control data analysis with real-time contextual information into actionable threat intelligence, the model uses an end-to-end pipeline for stream-based anomaly detection with three phases: behaviour prediction, inference and interpretation. Our earlier work [4] outlines the concept, while this work describes the technical realization and presents experimental results.

The remainder of the paper is organized as follows. Section II explains basic concepts and discusses related work, while Section III defines the technical problem. Next, Section IV describes the methodological development of the algorithmic framework and explains the core model of GenAttackTracker. Section V presents the experimental setup and the resulting insights, and Section VI concludes the paper.

II. Background and Related Work

Automation enables stable operation of OT: the devices and the machinery that monitor and control physical processes [2][3]; it enhances efficiency, quality of service delivery, productivity and safe operation of critical assets. Supervisory control of the cyber-physical system status is critical for issuing alerts and initiating an emergency shutdown operation when abnormal behavior patterns approach or violate defined safety margins.

A. Online Anomaly Detection

Supervisory control data is time series data to be interpreted as streamed real-value measurements taken at regular time intervals. Discordant patterns that do not match the expected normal system behavior but appear to occur "out of place" are called anomalies or outliers. Online detection of anomalous behavior in time series data streamed from the operation of an OT infrastructure can be a very challenging problem:

- Identifying anomalous behavior patterns requires learning normal behavior to train a robust machine learning model that not only fits previously observed data but also carries over to unobserved data. Developing such a model is usually not a trivial task.
- Anomalous patterns generally occur for various reasons, such as equipment failures, manual control intervention, and unauthorized tampering with control settings. Thus, an even more intricate problem is to differentiate the typically few anomalies of interest—above all, suspicious abnormal behavior indicating a potential security threat—from the vast majority of anomalies caused by noise, seasonality or other trends that are irrelevant to security.

Real-world physical processes are notoriously liable to difficult to predict "external" factors, such as fluctuations in demand and supply, technical instabilities, component failures et cetera. These phenomena result in hard to predict variance in the data—commonly referred to as "noise".

B. Suspicious Activity Markers

Indicators of Compromise (IOCs) are traditionally used in digital forensics to identify artifacts left behind by attackers, such as malware signatures, unusual traffic patterns, or file hashes. These are crucial in post-incident investigations, helping to trace and understand the extent of a breach [5][6]. IOCs are typically reactive, meaning they are often identified postmortem after the damage has already occurred [7].

In contrast, we present Suspicious Activity Markers (SAMs) as a concept aiming at real-time detection of threat activity before a cybersecurity compromise fully manifests. SAMs are akin to Indicators of Attack (IOAs), which have been promoted in industry contexts—originally by CrowdStrike—but with a distinct emphasis. IOAs generally focus on recognizing the Tactics, Techniques, and Procedures (TTPs) used by attackers. These indicators aim to detect cyberattacks at an early stage, potentially before significant harm is done and it is observable before the attack is fully unfolded. An IOA security strategy focuses on detecting the attacker's intent, enabling early intervention. Such indicators can assist security teams in intercepting even unknown types of attacks [7]–[9]. However, the definition of IOAs is vague and overlaps with IOCs, leading to potential confusion [10]. Examples of IOAs include but are not limited to [7]:

- Communication between public servers and internal hosts, indicating possible unauthorized data transfer;
- Connections through non-standard ports;
- User logins from multiple locations, potentially indicating stolen credentials.
- Unusual spikes in SMTP traffic;
- Internal hosts communicating with countries the business does not serve;
- Numerous honeytoken alerts from a single host.

Our specific focus here is on using SAMs as a secondary data source to corroborate findings from the primary source, i.e., supervisory control data. We define SAMs as follows:

Definition of SAM: A SAM is a contextual observation that provides additional insight into the operational security status of an SCADA system. SAMs are not intended to identify an attacker's intent directly, but rather to refine the understanding of potentially anomalous activities detected in supervisory

control data. By integrating SAMs with the primary data source, we aim to reduce false positives and improve the accuracy of detecting anomalies of interest, i.e., cyber threats to mission-critical infrastructures.

In previous works like [11]–[15], the integration of secondary auxiliary metrics into anomaly detection frameworks has been explored. Our approach differs significantly though in terms of generalization and method integration. Our focus is on using SAMs as secondary data sources to dynamically update our belief systems. This approach allows for a more refined and contextually aware detection mechanism.

C. Bayesian Analysis

Integrating contextual information from multiple sources can improve the effectiveness of cyberattack detection [16]. Bayesian modeling offers effective solutions for security threat detection and information fusion under uncertainty [4]. These methods can integrate heterogeneous data sources, including sensor networks and soft information, to improve anomaly detection in cybersecurity [17][18]. Bayesian models are particularly suitable for cyberattack detection due to their ability to update probabilities as new evidence is observed in addition to incorporating uncertainty. The continuous updating process makes Bayesian analysis and inference ideal for dynamic environments where attack patterns evolve over time [17]. Hierarchical Bayesian models, in particular, are well-suited for this context as they allow for multi-level aggregation of information, which is crucial for systems with distributed components and diverse data sources [19].

D. AttackTracker Framework

Attack Tracker is a distributed analytic framework designed for real-time detection of cyber threat activities in supervisory control system data [20][21]. By employing a scalable hierarchical network of detector agents to monitor various levels of the control system, the orchestration of threat detectors naturally matches the organization of SCADA architectures. Local detectors focus on identifying anomalies within subsystems, while higher-level detectors aggregate this information to detect and assess threats in different system components.

The framework consists of several key components. The Behavior Predictor learns and predicts normal subsystem behavior to identify any deviations or anomalies. The Inference Engine processes observations, assigns attack scores, and aggregates results from lower-level detectors to enhance system-wide threat detection. The Dynamic Scoring method adjusts detection thresholds dynamically based on the current system state and historical data, effectively handling contextual noise and reducing false positives.

Attack Tracker has been successfully applied to the Secure Water Treatment (SWaT) testbed [22] (see also Sect. V-A), demonstrating its capability to detect a wide range of cyber threats in SCADA-based CI systems.

III. Problem Definition

This section defines the problem of identifying anomalous behavior linked to a cyberattack in the supervisory control data streamed from the operation of a SCADA-based OT system. Henceforth, SCADA data is simply referred to as control data.

A. Primary and Secondary Data Sources

We categorize available data and information sources into a primary source and multiple secondary sources:

- **Primary Source:** Control data, formally represented as a multivariate time series $X = (x_t)_{t=1}^{T}$, for $T \in \mathbb{N}$, consists of discrete multivariate measurements x_t from sensors and actuators monitoring and controlling the system at time t, where t refers to logical rather than physical time. This data is the foremost anomaly detection input, providing insights into the operational state of the infrastructure.
- **Secondary Sources:** A given collection of SAMs is characterized as a set of 3-tuples, $\text{SAM} = \{\textit{SAM}_j\}_{j=1}^{N}$, where each $\textit{SAM}_j$ has three attributes, $(type_j, p_j, weight_j)$. Here, $type_j$ denotes the type of suspicious activity; $weight_j$ indicates the importance or impact of $\textit{SAM}_j$; and p_j represents the probability value indicating the likelihood of an attack being in progress. SAMs provide contextual information that can enhance the detection capabilities by highlighting potential threat indicators.

In order to effectively utilize both primary and secondary sources, we must establish a systematic approach that integrates multiple data streams, allowing for a comprehensive assessment of potential threats within the SCADA system.

At any time step $i > l$, with $i, l \in \mathbb{N}$, the supervisory control data to be analyzed at time i is given by X_i, for $X_i = (x_{i-l}, \ldots, x_i)$, while the corresponding activity marker values to be considered at time i are given by SAM_i, with $SAM_i = \{(\text{type}_{i,j}, p_{i,j}, weight_{i,j})\}_{j=1}^{N}$. The invariable length of the sliding observation time window is $l + 1$.

Objectives:

- Calculate the Anomaly Score AS_i at step i for $\mathbf{X}_i$, relative to the estimated behavior $\hat{\mathbf{X}}_i$, to assess any deviations of the actually observed from the expected normal behavior:

$$\text{AS}_i = f(X_i, \hat{X}_i), \text{with } f : \mathbf{X} \times \mathbf{X} \mapsto \mathbb{R}^+,$$

  where the real-valued function f quantifies the result.

- Update the posterior probability of an attack in progress, given the observed supervisory control data and the values of contextual markers (SAMs) at timestep i:

$$P(\text{Attack}_i | X_i, \text{SAM}_i) = \frac{P(X_i|\text{Attack}_i) \cdot \sum_{j=1}^{N} (p_{i,j} \times weight_{i,j}) \cdot P(\text{Attack}_i)}{P(X_i) \cdot P(\text{SAM}_i)} \quad (1)$$

  where:
  - $P(X_i|Attack_i)$ is the likelihood of observing the control data given an attack, informed by the Anomaly Scores AS_i,
  - $P(Attack_i)$ is the prior probability of an attack,
  - $P(X_i)$ and $P(SAM_i)$ are the marginal probabilities of the control data and SAMs, respectively.

The challenge lies in effectively integrating control data and additional contextual information to provide a comprehensive and real-time assessment of the threat level. A key difficulty is determining the extent of deviation from normal behavior that should be considered indicative of an attack, rather than a benign anomaly. This challenge of selecting the appropriate threshold for deviation is critical, as setting it too low may result in false positives (incorrectly identifying normal behavior as an attack), while setting it too high could lead to missing true positives (failing to detect actual attacks). Hence, developing a methodology that accurately analyzes these deviations and updates the likelihood of an attack based on new data and contextual markers is essential for more reliable threat detection. Figure 1 provides an overview of such a methodological framework. It depicts how primary supervisory control data X_i are processed by a machine learning model to generate an anomaly score. This score is compared against a threshold, with secondary data SAM_i integrated via a Bayesian model to refine the final attack likelihood score. The Bayesian analysis and inference are explored in detail in Sections IV and V.

B. Levels of Abstraction

To ensure a clear and structured approach, we consider different levels of abstraction in our problem definition:

- For the primary source (control data), we focus on detailed technical aspects, analyzing the data to compute Anomaly Scores AS_i. These scores reflect deviations between observed system behavior and the predicted normal behavior at timestep i. In this context, AS_i is used to assess the likelihood of the observed control data under different scenarios (attack vs. no attack).
- For the secondary sources (SAMs), we adopt a higher level of abstraction, where SAMs and their associated probabilities are assumed to be derived from external sources or specialized tools, which are integrated into our framework via APIs or similar interfaces. This approach allows us to efficiently incorporate diverse and potentially complex information into the decision-making process.

IV. Methodological Development of the Analytical Framework

While the dynamic scoring system of AttackTracker [20] achieves effective real-time anomaly detection on the SWaT testbed (see Sect. V-A), integration into a more comprehensive model enhances the inference process. Specifically, adding a hierarchical Bayesian module to the Inference Engine component broadens the scope of situational awareness to help reducing the rate of false positives. This extension allows for the inclusion of multiple secondary data sources and a more accurate assessment of the likelihood of cyberattacks.

A. GenAttackTracker

Our extension of the original AttackTracker model results in a generic model, named GenAttackTracker, which integrates dynamic anomaly scoring with a hierarchical Bayesian model. This dual approach leads to a more robust model for real-time anomaly detection by providing broader and deeper situational awareness for decision-making.

The main purpose of the hierarchical Bayesian model within the GenAttackTracker framework is to enhance the accuracy and reliability of cyberattack detection by integrating multiple sources of data, such as control data $\mathbf{X}_i$ and Suspicious Activity Markers (SAMs). The model continuously updates inferred beliefs about the likelihood of an attack in the light of new information becoming available.

1) Local Detectors: At the local detector level (Level 1), each detector monitors control data $\mathbf{X}_i$ to detect anomalies. Anomaly scores (AS_i) are computed here using modified z-scores [20], where the modified z-score (Z_s) is calculated as:

$$Z_s = \frac{X_i - \text{median}(X_i)}{\text{MAD}(X_i)}, \tag{2}$$

with X_i representing the observed data point at time i, and MAD is the median absolute deviation. The anomaly score AS_i at time i is defined as:

$$\text{AS}_i = |Z_s| \tag{3}$$

This score indicates the degree of deviation from expected behavior.

The prior distribution, representing the initial belief about the likelihood of an anomaly being an attack, is modeled based on historical SCADA data and the distribution of anomalies. Let θ_i^1 represent the prior belief at the local detector level:

$$\theta_i^1 \sim \text{Normal}(\mu_H, \sigma_H^2) \tag{4}$$

where μ_H and σ_H^2 are the mean and variance derived from historical SCADA data anomalies.

Bayesian inference is then applied to update these prior beliefs with new data, including the current SCADA data $\mathbf{X}_i$ and SAMs. The likelihood function incorporates the anomaly score AS_i and the SAMs, and modifies the prior distribution θ_i^1 to form the posterior distribution:

$$P(\text{Attack}_i|\mathbf{X}_i, \text{SAM}_i) \propto P(\mathbf{X}_i, \text{SAM}_i|\text{Attack}_i) \cdot P(\text{Attack}_i|\theta_i^1)$$

Expanding the likelihood function:

$$\begin{aligned} P(\mathbf{X}_i, \text{SAM}_i|\text{Attack}_i) &= P(\mathbf{X}_i|\text{Attack}_i) \cdot P(\text{SAM}_i|\text{Attack}_i) \\ &= P(\mathbf{X}_i|\text{Attack}_i) \cdot \prod_{j=1}^{N} (p_{i,j} \times weight_{i,j}) \end{aligned} \tag{5}$$

Plugging in the likelihood defined in Equation 5 and the prior $P(\text{Attack}_i|\theta_i^1)$, we obtain the posterior in Equation 6 as:

$$P(\text{Attack}_i|\mathbf{X}_i, \text{SAM}_i) = \frac{P(\mathbf{X}_i|\text{Attack}_i) \times \prod_{j=1}^{N} (p_{i,j} \times weight_{i,j}) \times P(\text{Attack}_i|\theta_i^1)}{P(\mathbf{X}_i, \text{SAM}_i)} \tag{6}$$

where $P(\mathbf{X}_i, \text{SAM}_i)$ is the marginal likelihood, ensuring that the posterior distribution sums up to one. The prior $P(\text{Attack}_i|\theta_i^1)$ reflects the initial belief about the likelihood of an attack, influenced by θ_i^1.

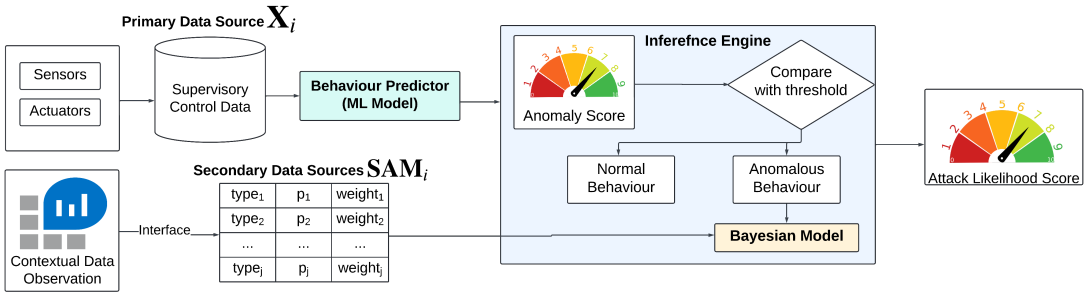

Figure 1. Integration of primary supervisory control data and secondary contextual data (SAMs) to compute anomaly and attack likelihood scores through a combined machine learning and Bayesian model.

2) Intermediate Levels: At the intermediate levels (Level $l \geq 2$), the information from multiple local detectors at the lower level $l-1$ is aggregated to refine the estimate of the attack likelihood at timestep i.

Each intermediate level l begins with a prior belief $\theta_i^{(l)}$, informed by the posteriors from Level $l-1$ as follow:

$$\text{Prior}_i^{(l)} = P(\text{Attack}_i^{(l)}|\theta_i^{(l)}) \Rightarrow$$
$$\text{Prior}_i^{(l)} = f\left(\{P(\text{Attack}_i^{(l-1)}|\mathbf{X}_i^{(l-1,k)}, \text{SAM}_i^{(l-1,k)})\}_{k=1}^{N}\right) \quad (7)$$

Equation 8 represents the likelihood aggregated from the underlying detector k at timestep i from Level $l-1$. N is the number of detectors contributing to the detector at the intermediate level l, and $w^{(l-1,k)}$ is the weight for the amount of contribution of each lower detector. In addition, we this weight normalized such that the sum of all weights ensures that the combined likelihood remains a valid probability.

$$\text{Likelihood}_i^{(l)} =$$
$$\sum_{k=1}^{N} w^{(l-1,k)}[P(\mathbf{X}_i^{(l-1,k)}|\text{Attack}_i^{(l)}) \cdot P(\text{SAM}_i^{(l-1,k)}|\text{Attack}_i^{(l)})] \quad (8)$$

The aggregated posterior at each detector in the intermediate level l is then computed as:

$$P(\text{Attack}_i^{(l)}|\mathbf{X}_i^{(l)}, \text{SAM}_i^{(l)}) = \frac{Prior_i^{(l)} \times Likelihood_i^{(l)}}{P(\mathbf{X}_i^{(l)}, \text{SAM}_i^{(l)})} \quad (9)$$

3) Global Detector: At the global level, the final assessment of the system's security status is made by aggregating information from the immediately preceding intermediate level L. The global detector can be seen as the final detector in the hierarchical structure, where it integrates all the aggregated information from the last intermediate level. In the following equations, (g) is the short form of global.

The prior distribution at the global level, denoted as $\theta_i^{(\text{global})}$, or $\theta_i^{(\text{g})}$ for short, is informed by the posteriors from the last intermediate level L at timestep i. This prior is formulated as:

$$\text{Prior}_i^{(\text{g})} = P(\text{Attack}_i^{(\text{g})}|\theta_i^{(g)}) \quad (10)$$

The likelihood at the global level is derived from the aggregated likelihood from the last intermediate level L, which has already integrated all the information from lower levels. $P(\mathbf{X}_i^{(L,k)}|\text{Attack}_i^{(\text{global})})$ is the likelihood of the SCADA data from detectors contributing to the global level at time i. The likelihood is expressed as:

$$\text{Likelihood}_i^{(\text{g})} =$$
$$\sum_{k=1}^{N_L} w^{(L,k)} \left[P(\mathbf{X}_i^{(L,k)}|\text{Attack}_i^{(\text{g})}) \times P(\text{SAM}_i^{(L,k)}|\text{Attack}_i^{(\text{g})})\right] =$$
$$\sum_{k=1}^{N_L} w^{(L,k)} \left[P(\mathbf{X}_i^{(L,k)}|\text{Attack}_i^{(\text{g})}) \times \prod_{j=1}^{N} (p_{i,j}^{L,k} \times weight_{i,j}^{L,k})\right] \quad (11)$$

Here, $w^{(L,k)}$ represents the weight assigned to the contribution of each detector k from the last intermediate level L.

The posterior probability at the global level at timestep i is then computed by combining the prior from Equation 10 and the likelihood from Equation 11:

$$P(\text{Attack}_i^{(\text{g})}|\mathbf{X}_i^{(\text{g})}, \text{SAM}_i^{(\text{g})}) = \frac{\text{Prior}_i^{(\text{g})} \times \text{Likelihood}_i^{(\text{g})}}{P(\mathbf{X}_i^{(\text{g})}, \text{SAM}_i^{(\text{g})})} \quad (12)$$

where:

- $\mathbf{X}_i^{(\text{g})}$ represents the aggregated SCADA data relevant to the global level at timestep i.
- $\text{SAM}_i^{(\text{g})}$ includes all SAMs to the global level at time i.

This approach ensures that the global level threat assessment integrates all available evidence at the current time step, taking into account the data and SAMs from all intermediate levels in the hierarchy. By continually updating the posterior probability $P(\text{Attack}_i^{(\text{global})})$ at each time step, the system maintains a comprehensive and accurate evaluation of potential threats, even when different detectors process different portions of the data.

Promoting a structured and methodical approach based on a hierarchical Bayesian model, GenAttackTracker integrates control data and SAMs at each time step i, thereby enhancing the detection and assessment of cyber threat activity. The result is a robust tool for improving situational awareness and decision-making in real-time.

V. EXPERIMENTS

In this section, we evaluate the performance of GenAttack-Tracker against the baseline AttackTracker framework using the SWaT dataset. The experiments aim to demonstrate the effectiveness of the enhanced dynamic scoring mechanism combined with Bayesian inference for real-time SCADA-based cyber-threat detection. We use Monte Carlo simulation for abstractly modeling externally determined SAM values as secondary inputs in the calculation of the posterior distribution.

A. Dataset

The SWaT dataset is derived from a Secure Water Treatment (SWaT) testbed, a scaled-down water treatment facility that simulates the operations of a real-world critical infrastructure system [22]. The dataset includes 11 days of continuous data, with the first seven days representing normal operations and the last four days containing multiple attack scenarios.

The dataset comprises 51 variables, including sensor readings (e.g., flow rates, water levels, pressure) and actuator states (e.g., pump statuses, valve positions), recorded at 1-second intervals. The result is a high-dimensional multivariate time series that serves as the basis for our analysis. Among these variables, the most critical for anomaly detection include, FIT201 (Flow Indicator Transmitter), LIT101 (Level Indicator Transmitter), PIT501 (Pressure Indicator Transmitter) and AIT502 (Analyzer Indicator Transmitter). These variables are particularly important due to their direct influence on the operational state of the water treatment process, making them key indicators of potential anomalies.

The dataset includes 36 distinct attack scenarios spread across the last four days, ranging from single-point disruptions to coordinated attacks affecting multiple components simultaneously. These scenarios are designed to simulate various real-world TTPs, such as tampering with sensor readings, manipulating actuator states, and disrupting communication between control components.

B. Implementation

The implementation of the GenAttackTracker framework was carried out in a Python-based tool environment, leveraging widely adopted libraries, such as TensorFlow for deep learning and Scikit-learn for statistical modeling. We used the PyMC3 library to implement the Bayesian inference process, allowing for efficient posterior estimation using Markov Chain Monte Carlo (MCMC) sampling. The implementation follows the steps outlined in Figure 2. For the experiments we used an Apple M1 Max chipset, featuring a 10-core CPU (3.2 GHz) and 32-core GPU, with 64GB of unified memory shared between CPU and GPU.

C. Data Analysis

In this analysis, we demonstrate how the hierarchical Bayesian model enhances and provides an experimental tool to study the effect of secondary data updating the probability of an attack with new observations. Figure 3 shows the combined likelihoods from four SCADA variables and SAMs. The spike

Input: SCADA data X, Suspicious Activity Markers (SAMs) S, anomaly score A
Output: Posterior probability of attack
procedure COMPUTELIKELIHOOD(X, A)
    Compute likelihood L based on SCADA data and anomaly score
    return L
end procedure
procedure CHOOSEPRIORS
    Set prior P_{attack} based on historical SCADA data
    Set prior P_{SAM} from external tools for SAMs
    return P_{attack}, P_{SAM}
end procedure
procedure UPDATEPOSTERIOR(L, P_{attack}, P_{SAM})
    Update posterior $P_{posterior} \leftarrow \frac{L \times P_{attack} \times P_{SAM}}{marginal\_likelihood}$
    return $P_{posterior}$
end procedure
procedure BAYESIANINFERENCE(X, S, A)
    $L \leftarrow$ COMPUTELIKELIHOOD(X, A)
    $P_{attack}, P_{SAM} \leftarrow$ CHOOSEPRIORS
    $P_{posterior} \leftarrow$ UPDATEPOSTERIOR(L, P_{attack}, P_{SAM})
    return $P_{posterior}$
end procedure

Figure 2. Bayesian Inference Engine Algorithm in GenAttackTracker.

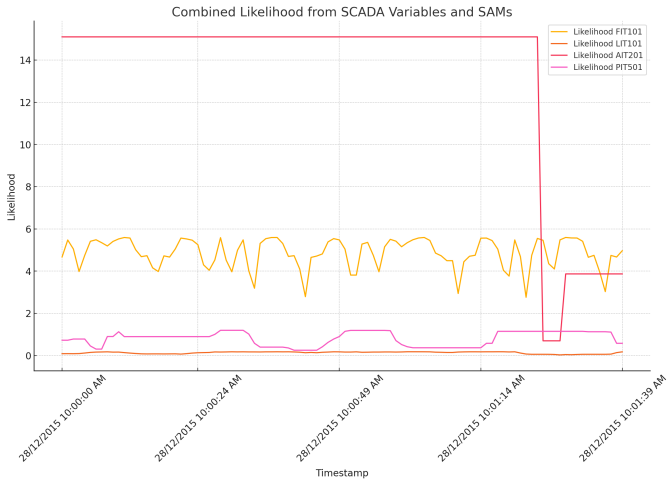

Figure 3. Combined Likelihood from four variables.

in the likelihood of LIT101 around 10:01:14 AM suggests a potential anomaly, possibly indicating an attack.

Figure 4 illustrates the evolution from prior belief to posterior distribution as new data is incorporated. Initially, the prior distribution reflects a low probability of an attack. As the anomaly in LIT101 and relevant SAMs are observed, the first posterior distribution shifts rightward, indicating an increased belief in the likelihood of an attack. A second posterior update further refines this belief, sharply increasing the probability and reducing uncertainty. These updates, informed by SAMs (summarized in Table I, demonstrate how integrating additional contextual data can enhance decision-making. The tighter confidence intervals in the global posterior

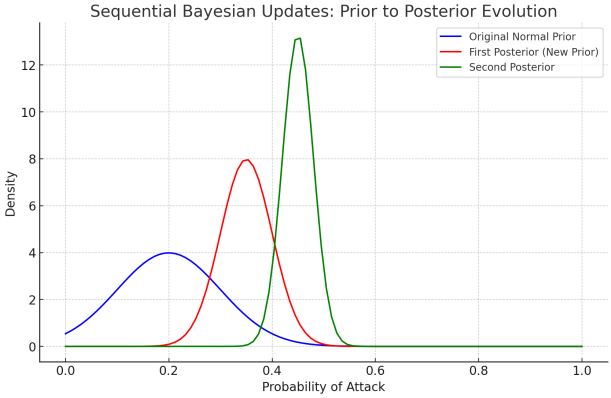

Figure 4. Updating the prior belief about the system security status.

TABLE I
SUMMARY OF SUSPICIOUS ACTIVITY MARKERS (SAMS)

SAM	Type	Probability (p_i)	Weight ($weight_i$)
SAM1	type_1	0.7	0.30
SAM2	type_2	0.6	0.20
SAM3	type_3	0.8	0.25
SAM4	type_4	0.5	0.15
SAM5	type_5	0.9	0.10

reflect a higher certainty in detecting actual attacks. The reduced variance demonstrates that GenAttackTracker can more confidently assess security threats in real-time by incorporating SAMs as secondary source of data.

VI. CONCLUSION

In this paper, we introduce GenAttackTracker, an innovative framework for enhancing real-time detection of cyber threats targeting SCADA-based critical infrastructure systems. By integrating dynamic anomaly scoring with hierarchical Bayesian models, GenAttackTracker addresses the complexities of identifying and interpreting cyber threats within highly interconnected operational technology environments. A key contribution of this framework is its ability to incorporate SAMs as secondary contextual data, providing a more assured threat detection process. This integration not only reduces the likelihood of false positives but also allows the framework to serve as a powerful experimental tool by evaluating the effects of secondary input data on the overall system status, using Monto Carlo simulation in the calculation of posterior distributions. By doing so, it enhances decision-making processes and improves situational awareness. The experimental results confirm that the inclusion of contextual information refines threat assessment, making this approach a valuable addition to the cybersecurity domain. While putting a spotlight on SCADA, the strategies we discuss here do likely apply to a much broader range of industrial process control systems.

In our continued work, we plan to further generalize the GenAttackTracker model, going beyond analyzing isolated OT infrastructures, to analyze cyber threat activities across ecosystems of linked critical infrastructures as outlined in [4].

REFERENCES

[1] O. S. Saydjari, "Engineering trustworthy systems: a principled approach to cybersecurity," *Communications of the ACM*, vol. 62, no. 6, pp. 63–69, May 2019. [Online]. Available: https://dl.acm.org/doi/10.1145/3282487

[2] K. Stouffer *et al.*, "Guide to operational technology (ot) security - nist sp 800-82r3," *NIST Special Publication*, pp. 800–82, September 2023.

[3] Fortinet, "What is OT Security?" Online: https://bit.ly/3XkP0Bk, 2024, accessed: 2024.08.31.

[4] F. Movafagh and U. Glässer, "Cyber situational awareness of critical infrastructure security threats," in *The Eighth International Conference on Cyber-Technologies and Cyber-Systems, CYBER 2023, Porto, Portugal, Sept./Oct., 2023*. IARIA, 2023, pp. 53–61.

[5] M. Asiri, N. Saxena, and P. Burnap, "Investigating usable indicators against Cyber-Attacks in industrial control systems." USENIX Association, Aug 2021.

[6] M. Asiri, N. Saxena, R. Gjomemo, and P. Burnap, "Understanding indicators of compromise against cyber-attacks in industrial control systems: a security perspective," *ACM transactions on cyber-physical systems*, vol. 7, no. 2, pp. 1–33, 2023.

[7] Sai, "Indicator of Compromise (IoC) vs. Indicator of Attack (IoA)," Online: https://bit.ly/4dDu1PR, August 2022, accessed: 2024.8.31.

[8] Muhammad Raza, "What Are IOAs? Indicators of Attack Explained," Online: https://splk.it/3Xo26hh, May 2023, accessed: 2024.08.31.

[9] CrowdStrike, "IOA VS IOC," Online: https://bit.ly/3MpxaGU, October 2022, accessed: 2023.08.31.

[10] E. Kost, "What are IOAs? How they differ from IOCs," Online: https://bit.ly/4g27UnQ, April 2023, accessed: 2024.8.31.

[11] M. Almgren, U. Lindqvist, and E. Jonsson, "A multi-sensor model to improve automated attack detection," in *Recent Advances in Intrusion Detection: 11th International Symposium, RAID 2008, Cambridge, MA, USA, September 15-17, 2008. Proceedings 11*. Springer, 2008, pp. 291–310.

[12] M. J. Pappaterra and F. Flammini, "A review of intelligent cybersecurity with bayesian networks," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE, 2019, pp. 445–452.

[13] D. Lin, A. Li, and R. Foltz, "Beam: An anomaly-based threat detection system for enterprise multi-domain data," in *2020 IEEE International Conference on Big Data (Big Data)*. IEEE, 2020, pp. 2610–2618.

[14] T. Bass, "Intrusion detection systems and multisensor data fusion," *Communications of the ACM*, vol. 43, no. 4, pp. 99–105, 2000.

[15] N. Bakalos *et al.*, "Protecting water infrastructure from cyber and physical threats: Using multimodal data fusion and adaptive deep learning to monitor critical systems," *IEEE Signal Processing Magazine*, vol. 36, no. 2, pp. 36–48, 2019.

[16] A. AlEroud and G. Karabatis, "Beyond data: Contextual information fusion for cyber security analytics," in *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, 2016, pp. 73–79.

[17] J. A. Perusquía, J. E. Griffin, and C. Villa, "Bayesian models applied to cyber security anomaly detection problems," *International Statistical Review*, vol. 90, no. 1, pp. 78–99, 2022.

[18] K. Wu, W. Tang, K. Z. Mao, G.-W. Ng, and L. O. Mak, "Semantic-level fusion of heterogenous sensor network and other sources based on bayesian network," in *17th International Conference on Information Fusion (FUSION)*. IEEE, 2014, pp. 1–7.

[19] A. Gelman *et al.*, *Bayesian Data Analysis (3rd Edition)*. CRC Press, 2014.

[20] Z. Zohrevand and U. Glässer, "Dynamic attack scoring using distributed local detectors," in *ICASSP 2020-IEEE International Conference on Acoustics, Speech and Signal Processing*, 2020, pp. 2892–2896.

[21] Z. Zohrevand, "End-to-end anomaly detection in stream data," Ph.D. dissertation, School of Computing Science, Simon Fraser University, 2021.

[22] J. Goh, S. Adepu, K. N. Junejo, and A. Mathur, "A dataset to support research in the design of secure water treatment systems," in *Critical Information Infrastructures Security: 11th International Conference, CRITIS 2016, Paris, France, October 10–12, 2016, Revised Selected Papers 11*. Springer, 2017, pp. 88–99.