IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON EDGE INTELLIGENCE FOR INTERNET OF THINGS

Received May 6, 2020, accepted May 20, 2020, date of publication June 1, 2020, date of current version June 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2998983

Capturing-The-Invisible (CTI): Behavior-Based Attacks Recognition in IoT-Oriented Industrial Control Systems

AKASHDEEP BHARDWAJ[1], FADI AL-TURJMAN[2], (Member, IEEE), MANOJ KUMAR[1], THOMPSON STEPHAN[3], AND LEONARDO MOSTARDA[4], (Member, IEEE)

[1]School of Computer Science, University of Petroleum and Energy Studies, Dehradun 248007, India
[2]Research Center for AI and IoT, Artificial Intelligence Department, Near East University, Mersin 10, Turkey
[3]Department of Computer Science and Engineering, Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida 201313, India
[4]Computer Science Division, Camerino University, 62032 Camerno, Italy

Corresponding author: Thompson Stephan (thompsoncse@gmail.com)

ABSTRACT Industrial Control Systems monitor, automate, and operate complex infrastructure and processes that integrate into critical industrial sectors that affect our daily lives. With the advent of networking and automation, these systems have moved from being dedicated and independent to centralized corporate infrastructure. While this has facilitated the monitoring and overall management using traditional detection methods, Web Application Firewalls or Intrusion Detection Systems has exposed the networks subjecting them to Behavior-based cybersecurity attacks. Such attacks alter the control flow and processes and have the malicious ability to alter the functioning of these systems altogether. This research focuses on the use of process analytics to detect attacks in the industrial control infrastructure systems and compares the effectiveness of signature-based detection methods. The proposed work presents a pattern recognition algorithm aptly named as ''Capturing-the-Invisible (CTI)'' to find the hidden process in industrial control device logs and detect Behavior-based attacks being performed in real-time.

INDEX TERMS Industrial control systems, cyberattacks, behavior detection, signatures.

I. INTRODUCTION

Industrial Control Systems (ICS) are integrated infrastructures to control industrial systems distributed over large geographical areas and locations [1]. These include networks, sensor devices, and controllers to automate and operate industrial tasks and processes effectively. Industrial Control Systems are either Distributed Control Systems (DCS), Supervisory Control & Data Acquisition (SCADA), or Hybrid systems that combine the best features of both the models. These industrial systems are highly important to arranging critical infrastructure industrial processes that are core to our lives. Historically, manufacturing and engineering components are used in operations of generation, distribution, the transmission of water, energy, food, manufacturing, and other critical infrastructures [2]. Industrial Control Systems include management control and data acquisition (SCADA) systems, distributed control systems (DCS), and other control system configurations such as programmable logic controllers (PLC). ICSs are found in the industrial sectors and critical infrastructures, such as nuclear and thermal plants, water treatment facilities, power generation, heavy industries, and distribution systems. Any compromise to ICS leads to enormous physical and environmental damage as well as a danger to human lives. ICS uptime necessitates 100% availability, which in turn proves difficult and costly to pause or interrupt for maintenance or patching for security updates [3]. Since ICS relates to physical impacts, the impact of even small downtime can affect millions [4]. ICS comprises of the following critical components as shown in Figure 1.

Information Technologies (IT) & Operational Technology (OT) include critical software and hardware systems for the control and monitoring of physical sensor field devices. IT and OT provide essential, inherent integration and visibility for supply chain details about logistics, assets, processes, and completion times. This provides remote control and management units with information, thus keeping the ICS efficient and competitive. However, IT and OT are often

The associate editor coordinating the review of this manuscript and approving it for publication was Sherali Zeadally.

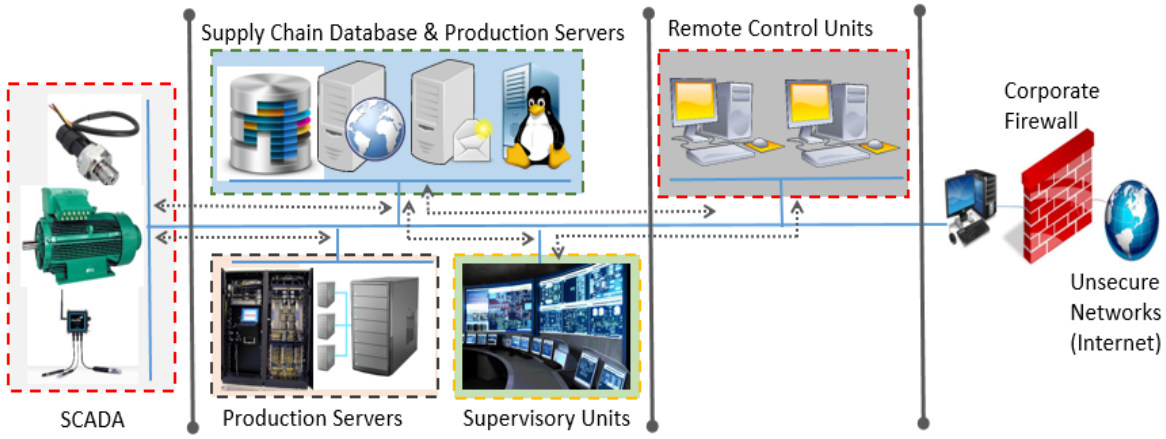

FIGURE 1. Industrial control systems components.

targeted by cyberattackers, as most of the ICS do not have stringent security policies or the infrastructure to detect and monitor cyberattacks [5].

- Human Machine Interface (HMI) provides a graphical user interface (GUI) application that assists the interaction of hardware, control system, human operators (staff). HMI displays trends, historical and real-time status from data and logs gathered from the ICS environment. MI provides the dashboards to monitor, customize, set control points, and establish the operational parameters required for the day-to-day sensor and controller [6].
- Programmable Logic Controller (PLC) is the control component of the ICS ad that provides process management. PLC provides supervisory, remote access, and control to devices such as actuators and sensors [7].
- Remote Terminal Units (RTU) & Master Controller Units (MTU) are microprocessor-based field devices. RTU receives commands from the MTU and sends back the information from the field.

Control Server &Loops host supervisory control systems, communicate with each low level, on-field control devices such as PLC and actuators to carry out tasks and complete processes. The control loop interprets sensor signals, motors, gears, control valves, breakers, and other electromechanical devices. Intelligent Electric Device (IED) are smart devices that acquire data, communicate with other devices to control and perform local processing automatically. Remote Maintenance & Diagnostics identifies and prevents abnormal operations or failures and helps to prevent hardware and software related problems inside ICS.

II. INTEGRATION OF IT AND OT

ICS Infrastructure and networks were initially air-gapped, and they remained secure from the outside world, majorly from the insecure Internet. However, this changed in the past few years as the applications and components, directly or indirectly communicated over the internet. Most ICS were set up several decades ago, the legacy systems and applications running for ICS were computerized that still worked on proprietary protocols and weak network designs. Advances in smart sensors, Internet of Things (IoT), and wireless networks integrated with the use of Operational Technology (OT) and Information Technology (IT) for leveraging the high speed, real-time response, and cost-effectiveness. The arrival of new technologies like Virtualization, Cloud Computing, Software Defined Networks, Big Data Analytics, IoT, Machine Learning, and Artificial Intelligence let to a huge improvement in industrial productivity and system functions. This required integration of OT and IT with the outside world, mainly using the Internet. This integration between these is shown in Figure 2.

ICS performs repetitive and restricted tasks, so under normal operations, the systems, devices, network, and sensors record a standardized set of parameters, logs, and processes. ICS logs compriseof highly sensitive and critical information that is analyzed to detect major variances and device disruptions in the task control flow or processes of the operational infrastructure. However, detection and monitoring of low-level data variance, process delays, or network probe scans are highly unlikely to be successful using the traditional signature-based security systems. These new cybersecurity attacks on ICS can disrupt the sequence of events, processes, and the control flow [8], [9].

Process Mining is used traditionally in business operations, some researchers proposed the use of process mining techniques to detect anomalies in the control flow of industrial control systems [10]. Business operations focus on the use of process mining to discover events, monitor non-conformance, and perform process improvements. This method uses logs generated by networks, systems, and devices, which have details for an event and regarding the activity conducted with a timestamp and an Event ID relating that specific process instance [11]. Since ICS generates tons of logs, the log traf-

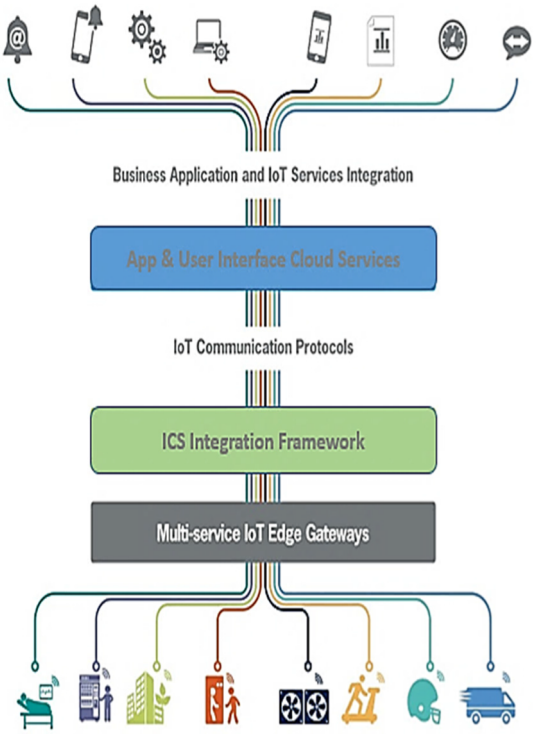

FIGURE 2. Integration with OT & IT.

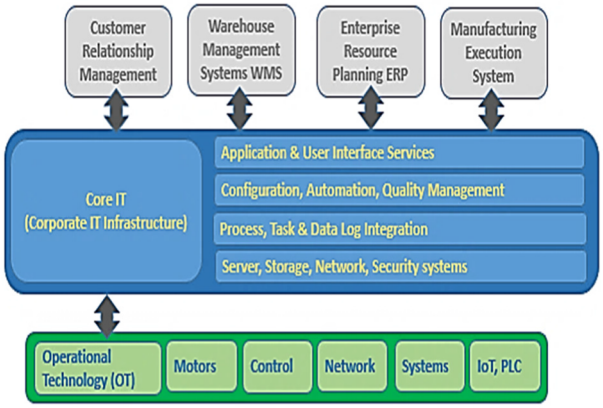

FIGURE 3. IT and OT Integration.

fic is routed to various remote monitors and management control units. As presented in Figure 3, these remote terminal units control the actual field devices, which include Sensors, Actuators, PLCs, Circuits, Motors, and Pumps.ICS infrastructures, integrate OT and IT and merge control and management of devices, operating systems, and software. This integration results in increased network capabilities and bandwidth, improved reliability of systems, and efficient log and data gathering response. However, the advantage also presents a huge surface area to target for the cyberattackers.

This is mainly due to vulnerabilities inherent in legacy operating systems, application, centralizing data, and old control systems models in SCADA as well as introduces previously unknown vulnerabilities. Advanced level attack scenarios like behavior-based attacks are monitored and detected for any malicious or suspicious pattern. These are flagged as alerts and sent for analysis. Traditional signature-based security tools monitor and identify only the pre-listed patterns and known threats and it does not detect unknown and unidentified suspicious threats. ICS is less secure against such behavior-based attacks.

III. LITERATURE SURVEY

This section widely discusses the related works on Industrial Control and Cybersecurity topics. Wang *et al.* [12] proposed a software-defined network (SDN) based dynamic cybersecurity protections for industrial control systems (ICSs). Their method involved executing real-time SDN security response measures such as redirection and isolation. This formed the security detection response control with moving target defense. This protected the ICS using SDN topology and port hopping to deceive and confuse the attackers. This prevents attacks and protects the ICS in active mode.

Soufian [13] proposed an alternate but practical solution for basic key security and common security threats such as Flooding attack and associated risks against industrial automation systems. The author implemented countermeasures directly at the endpoints irrespective of hardware devices or deployment platforms. Their solutions proposed the use of a control algorithm with self-defending control systems and defense mechanisms.With the increase in data networking in ICS, cybersecurity challenges have grown into a critical problem. The use of risk assessment can be vital for ICS protection. However, the risk propagation model is difficult to build due to the lack of historical datasets. In 2018, Zhang *et al.* [14] presented a Fuzzy Probability Bayesian Network for dynamic risk assessment. The approach initially established for prediction and analysis of the propagating cyber risks. To overcome the issue of limited historical data, the use of crisp probabilities is adopted and replaced with fuzzy probabilities. Experiments conducted on chemical reactor control systems successfully demonstrated the effectiveness of the proposed approach. Cybersecurity attacks on ICS, Industrial IoT, and Industrial 4.0 architectures and infrastructures are increasing every year. Threats to cybersecurity are mostly discovered after the breach. Lou *et al.* [15] proactively analyzed cybersecurity attacks from the system functionality. Apart from considering confidentiality and availability, the authors also focused functional and information integrity. During the analysis phase, this delivered an accurate determination of cybersecurity issues. The authors conducted a cybersecurity analysis of nuclear power ICS and presented the final analysis in casual fault graph and attack models that illustrated possible attack vectors from the analysis. Gómez *et al.* [16] generated reliable anomaly detection for datasets in Railways ICS. Their methodology consisted of attacks-selection, attack-deployment, traffic-capture, and feature-computation. The authors trained several Deep and Machine Learning models

to detect anomalies and trials to prove that their models have high precision and suitability for use in real-time production systems. Abdelghani [17] presented ICS Security such as supervisory control and data acquisition implemented in power transmission networks, stations, and distribution grids. The authors provided recommendations between ICS and IT security to avoid intrusion and destruction of industrials plants. Lou *et al.* [18] suggested practical and unique approaches for verifying the extensiveness and precision of functional specifications for complex safety ICS. The authors combined Artificial Intelligence planning techniques with the formal functional specification. This assisted better analysis of cybersecurity vulnerabilities.

ICS today face more cybersecurity issues than before, leading to all the more severe risks in critical infrastructure. To mitigate such critical risks, an appropriate security strategy needs to be developed. However, due to the lack of consideration of the strategy for securing physical and cyber domains, there remains a gap in the tradeoff between ICS requirements and security implementation. To overcome such limitations, Li *et al.* [19] presented a decision-making approach for intrusion response in ICS. They maximized the objective benefit vectors of security, system, and state. Then the closest solutions to the ideal security strategy were chosen and the efficiency of the proposed approach demonstrated high value in a simulated process control ICS. Escudero *et al.* [20] provided comprehensive outlines, vision, and views of ICS Security for the G-SCOP Center of Research deficiencies. They gave orientation for designing a behavioral model-based Intrusion Detection System for equipment degradation and addressed the limitation of the current approach on a single supply chain. Angle *et al.* [21] demonstrated historical examples of real damage to cyber-physical systems. The authors analyzed threats posed by software-controlled variable frequency drives (VFDs) and designed a prototype version of a simulated attack on ubiquitous equipment on ICS. Modern ICS faces a rising number of cybersecurity issues with the adoption and integration of information, communications, and network technologies which leads to severe risks to infrastructure and assets. Therefore, Li *et al.* [22] presented an approach for dynamic impact assessment. This approach predicted the impact trend for full asset recognition.

The role of cybersecurity assurance is highly critical for managing trust in smart grid communication systems. Ogundokun *et al.* [23] provided innovative risk-based insights for approaches and baseline to cybersecurity assurance in smart grid and automation systems. Cybersecurity assurance baselining is implemented as per security impact levels to manage the trust. They selected and justified security assurance controls by using the US Defense Information Systems Agency's Security Technical Implementation Guides for control and selection of national security systems. Leander *et al.* [24] discussed the main challenges faced in ICS and IoT systems concerning cybersecurity. They presented findings as a flow-control loop applied to a simple threat model and deduced cybersecurity requirements [25]. This process achieved flexible, efficient, and affordable cost-effective production for industrial automation of digital transformation in ICS [26], [27].

From the literature review, the key research gaps were identified and appropriate solutions are proposed to fill these research gaps, which makes the proposed work novel.

The identified research gaps are presented below:

- **Create ICS Process models for human interpretation**: These models can be used as an input for process mining activities, such as conformance checking in device logs and used with process discovery algorithms.
- **Application of security conformance checks**: Currently conformance checks are conducted within business contexts. This can range from reviewing tasks, security audits, network security, or baselining security for IoT and Sensors.
- **Selection of Process Discovery algorithms**: Use of Process Discovery can be customized for this research to detect duplicate tasks, incorrect event sequences, close looping processes, auto-allocate free resources, and even detect invisible or silent tasks. Another key area is identifying application issues and ensure the tasks and processes run without failure or the wait times, even if there is no alert generated.

Considering the above-mentioned research gaps, this paper focuses on the below-mentioned objectives:

- For human interpretation of the ICS Process models, there is an opportunity to determine the suitability for best-fit algorithms for the various ICS categories, each having different or some common features and capabilities.
- Conformance checks for ICS Security can be resolved by implementing and integrating the best practices for each system, network, and device into the ICS Security Policy. This should include mandatory device logging and recording, separation of duties for software modules, and aggregating the device logs to an event collector for process mining based analysis.
- The authors proposed a new process discovery algorithm that can detect and monitor issues from device logs without any Alert ID from PLC datasets. The proposed algorithm is named as 'Capturing-the-Invisible' algorithm.

IV. METHODOLOGY AND EXPERIMENTAL SETUP

ICS testbed experimental prototype was designed, as the conveyor belt model having PLC hardware, HMI, Raspberry PI, and infrared barrier sensors. Conveyer belt workflow is managed by the PLC with input sensors providing process information as logs for the PLC. The operational system resembled real conveyor and generated log data representing an industrial setup. To acquire and analyze logs for anomaly detection, Raspberry Pi is integrated into the testbed, which runs the cybersecurity attacks. This generated data attack traffic of anomaly detection with known and unknown classification, and even alerts. The overall workflow, events are controlled

TABLE 1. Process model sequence and log event activities.

Events occurring in the logs → T(Log) = {t ∈ T |∃σ∈W t ∈σ},
Start of log trace event activities → T(Start) = { t∈ T |∃σ∈ log t = first(σ) },
End of log trace event activities → T(End) = { t∈ T |∃σ∈ log t = last(σ) },
X(log) = { (A,B) | A ⊆ T(log) ∧ B ⊆ T(log) ∧∀a ∈A∀b∈ B a →log b ∧∀a1,a2 ∈ A a1#log a2 ∧∀b1,b2 ∈ B b1#Log b2 },
Y(log) = { (A,B) ∈ X |∀(A′,B′) ∈ XA ⊆ A′∧B ⊆ B′⇒ (A,B) = (A′,B′) },
P(log) = { p(A,B) | (A,B) ∈ Y(Log) } ∪{I(Log),o (Log)},
F(log) = { (a,p(A,B)) | (A,B) ∈ YW ∧ a ∈ A } ∪ { (p(A,B),b) | (A,B) ∈ Y(log) ∧ b ∈ B } ∪{ (I log,t) | t ∈ TI} ∪{ (t, o log) | t ∈ TO},
α(W) = {P(log), T(log), F(log)}.

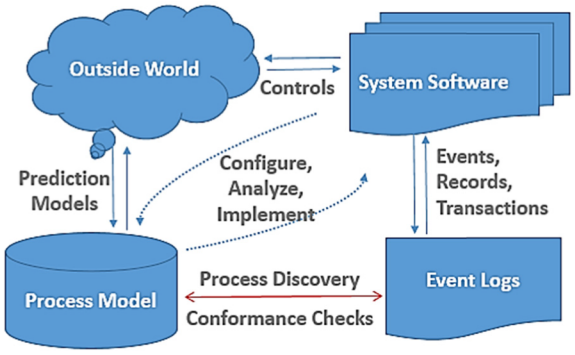

FIGURE 4. End to End workflow for industrial control systems.

and managed by Siemens PLC. For comparing the logs with existing models, the connected network, physical components of the experimental setup with sensors connected to the logic controllers, and the external cyber network are shown in Figure 4.

In this research work, cyberattacks were performed to create and record datasets from two industry-standard ICS PLC devices as mentioned below. The intent of performing attacks on both datasets was to disrupt or change the standard running industrial processes and tasks in the PLCs.

- The First dataset was generated from Siemens S7-1200, this dataset logged control data device logs, normal network traffic packets, processes, and attack data from cyberattacks. These cyberattacks involved in several unique attacks, which mainly include injection and flooding attacks.
- The Second dataset was generated from National Instruments NI-cRIO-9074, this dataset comprised device logs and network packet captures. This covered the normal behavior of devices and processes and attack traffic from cyberattacks.

To compare the CTI Algorithm with existing processes mining discovery approaches, the following investigations were proposed:

- Process models generated by the CTI algorithm were analyzed as per the modeling algorithm requirements.
- Compared and contrasted the process model generated by each algorithm, matching with the ICS modeling requirements identified from the previous work.
- Imported process models generated by the CTI algorithm into the process of mining toolkit and conducted process mining conformance checks. This compared the process model generated by the CTI algorithm with the models generated by the process discovery algorithms on the two datasets.
- Results of conformance checks are compared with existing process mining algorithms to determine if the CTI algorithm can build a new model that identifies anomalous events in ICS device logs.

For the process model, the sequence of events in the first iteration is (A, B, D) and for the second iteration process, the sequence of events is (A, C, D). This is represented in the Petri-net process model that maps the processes control flow. This process of modeling and discovering the sequence of events is applied to ICS tasks to discover hidden attacks and process models on the control flow of the ICS process. For implementing the process discovery algorithm, an Alpha algorithm was implemented. This starts with a set of activities T, event log (L) overthe event (A) →L∈B(A *), and (a, b∈A) as shown in Table 1.

A. CAPTURING-THE-INVISIBLE (CTI) ALGORITHM

The focus is to design the new algorithm to discover a process model that can be adapted for use in process mining based toolkit. Input for the CTI algorithm includes logs from the ICS devices and the output of the algorithm is the process model in Petri-net form. CTI algorithm depends on certain assumptions discussed below.

- It is assumed that the input is a complete device log that contains an ordered list of items in the order the items were recorded by the HMI and ICS devices. The sequence needs to be consistent.
- It is assumed that the log consists of at least one timestamp and one item. The timestamp is used for ordering.
- For two items with similar timestamps, the sequence is maintained in the same order in which the logs are recorded.

In this work, the proposed CTI algorithm is presented using various stages, each performing different tasks.

For the CTI algorithm, the new process mining anomaly is validated, and then the detection method identifies variances and anomalies including cyberattacks in two datasets. Logs are generated from industry-standard industrial control

TABLE 2. Comparing algorithms (anomalies detected, conformance & time taken).

Algorithms	Anomalies Discovered	Conformance %	Time Taken (ms)
Alpha	13	77%	18.5
Heuristic Miner	22	53%	28.4
ILP Miner	15	65%	21.7
Manual Method	26	48%	32.5
Genetic Miner	13	77%	17.2
Fuzzy Miner	21	55%	29.8
CTI Algorithm	34	93%	11.9

system devices, one from Siemens and other from National Instruments.

The device logs are transformed into event logs and similar events are paired. From these events, unique pairs are discovered and followed after each event. Further, events in the loop are discovered. Finally, silent and hidden transitions are discovered. Various stages of the proposed CTI algorithm are discussed below.

Stage 1 Algorithm to Transform ICS Device Logs into Event Logs

```
# Input Data: Device Log from ICS
# Output Result: Pre-processed Event Log for Process Mining
START
      initialize;
while!EOL
         # Not End of Device Log File
         do
            read log rows;
            for rows in device log
            do
# Remove noise data
Removeϵ Non-event Records();
            Remove ϵFiltered Attributes();
            # Add required Case Identifiers and
Human/Staff names
            Add ϵ Case-ID();
            Add ϵ Operator-Staff-Names();
            # Display output
print (“Non-events and Filtered attributes removed.”)
            print (“Added Case IDs and Staff Names.”)
            Collate ϵ Log();
      end
   end
  END
```

Five different stages for the proposed CTI algorithm is presented for the events describing behavior with inputs from the device logs for two datasets. Given our working example, the output of this stage is presented as a set of events, in the form of e = (iP, T, oP). An example can be, of the set L = A, B, C, D, E and the series of events ((p1, A, p2), (p2, B, p3), (p3, C, p4),(p4, D, p5), (p5, E, p6)).

Stage 2 Algorithm Module to Create Event Log Pairs

```
# Input Data: Device Log from ICS
# Output Result: Set of unique pairs
START
      A = set of activities (a));
      P = set of pairs (p);
      LI = last activity;
         P ϵ Ø;
         LI ϵ Ø;
      for all a 2 A do
(p) = {(LI, a) };
         LI = a;
end
Uniq(p) = set of all unique pairs (p);
Uniq(p) = Ø;
for all p ϵ P do
      if (p) !ϵ Uniq(p) then
Uniq(p) ≤ {p};
      else
         Pass;
      end
END
```

V. EXPERIMENTAL RESULTS

The process models are evaluated using the CTI algorithm and compared with models generated with extensively used process discovery algorithms. The CTI algorithm is evaluated using pre-identified core capabilities and requirements of process discovery algorithms. Validation of the process models discovered is conducted from the two generated datasets. The requirement included the ability to discover invisible or duplicate tasks, event sequences, loops, and the unlabeled

Stage 3 Algorithm Module to Discover Sequence Events

```
# Input Data: Uniq(p) is set of pairs
# Output Result: Marking Sets
START
(p) = {(a1, a2)}
M = set of all events e = (iP, T, oP);
S = set of all valid events;
C(P)= Counter for (iP, oP);
S = Ø;
C(P)= Ø;
for all p ϵ Uniq(P) do
 if a1 !ϵ S then
      if a2 !ϵ S then
      M U {(iP, a1, oP)};
      S U {a1};
      CP++ M {(iP, a2, oP)};
 S U {a2};
      CP++;
      else
         Pass;
   end
else
   Pass;
end
END
```

Stage 4 Algorithm Module to Discover Looping Event Logs

```
# Input Data: Uniq(p) where Uniq(p) is a set of pairs
(p) = {(a1, a2) }
# Output Result: Marking with loop set
START
M = set of events e = (iP, T, oP);
S = set of all seen events;
C(p) = counter for iP, oP;
S = Ø;
C(P)= Ø;
for all p ϵ UP do
    if a1 !ϵ S then
      if a2 ϵ S then
        for e ϵ M do
          if a2 U M then
             Pi = e(iP );
          end
        M U {(iP, a1, Pi)};
        CP++
        end
        else
          Pass;
        end
      else
        Pass;
    end
END
```

Stage 5 Algorithm Module to Discover Silent and Hidden Transitions

```
# Input Data: Uniq(P) where Uniq(P) is set of pairs
(p) = {(a1, a2)}
# Output Result: Silent and hidden transitions
START
M = set of events e = (iP, T, oP);
S = set of all seen events;
C(P) = counter for iP, oP;
C 0 counter for 0;
C= Ø;
   for all p ϵ Uniq(P) do
    if a1, a2 ϵ S then
     for all ϵ 2 M do
      if iP = oP then
        M U {(a1(oP), ρ∈, a2(oP))};
        CP++
      else
        Pass;
      end
     end
    else
     Pass;
end
END
```

event identifier. Two process models were identified when using our algorithm while none is found from other state-of-art algorithms. This indicated that the proposed algorithm discovers event sequences, unique item pairs, and events, as shown in Table 2. Process models having invisible tasks are discovered by the CTI algorithm. These silent tasks pass through the model in forwarding or backward pathways in first level check. The proposed process discovery algorithm discovered process models from device logs without the use of Case Identifier or Alert generation with using the Siemens PLC hardware. Figure 5 represents the graphical representation of obtained results which indicates that the CTI algorithm is producing better results.

Results displayed from Table 3 show that the CTI algorithm delivers better results for both true positive and false negative conformances with the same requirements and inputs for inductive miner algorithms. Results for this check are obtained by comparing the IM algorithm and the CTI algorithm for true and false conformance. A graphical representation for first level conformance can be seen from Figure 6 which depicts the second level of conformance outputs.

The second level of conformance check results using event log from National Instruments dataset with Inductive Miner algorithm is presented in Table 4. The same experiment is repeated with the process model generated by the

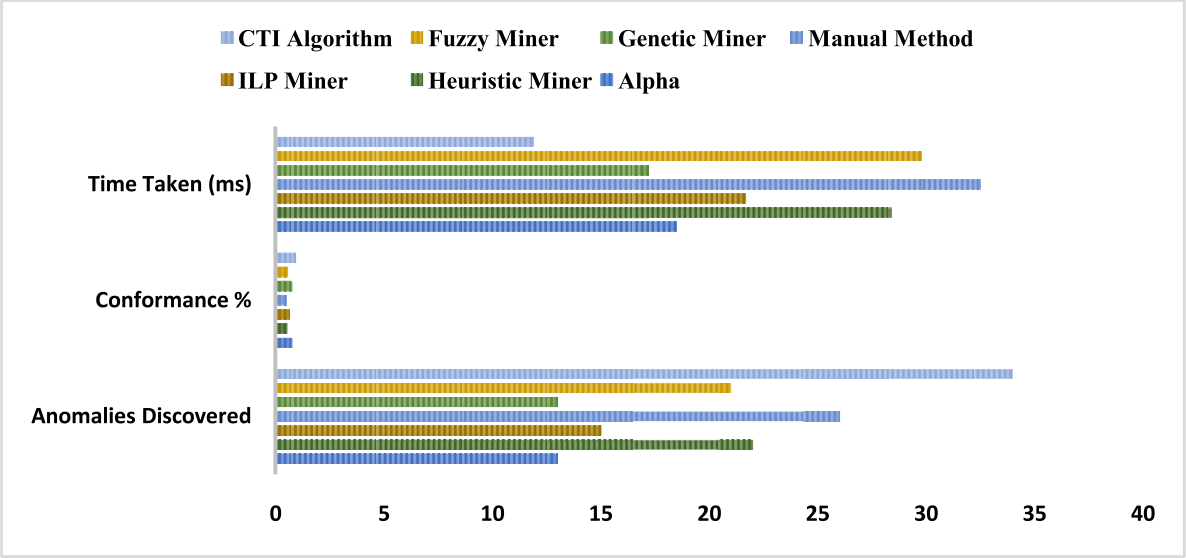

FIGURE 5. Performance comparison between CTI and other algorithms.

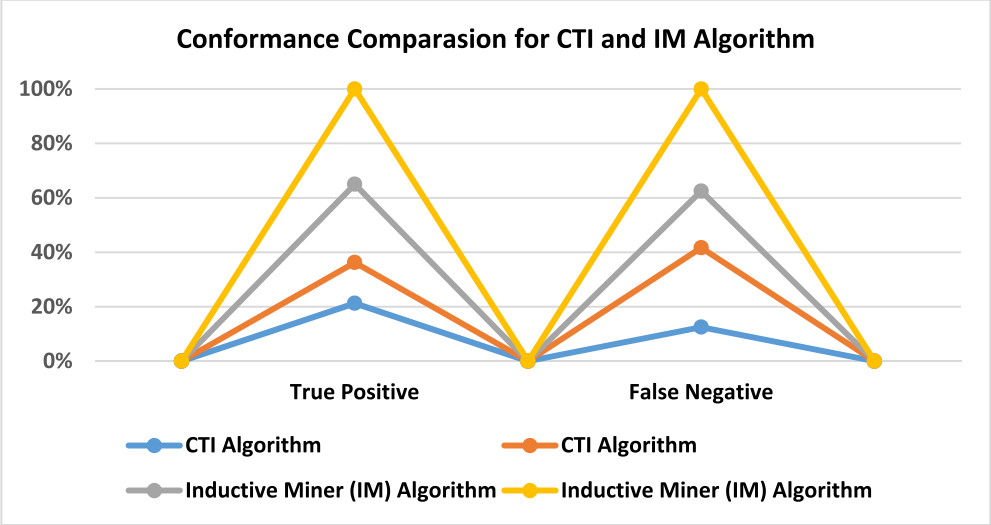

FIGURE 6. First level conformance comparison in terms of true positive and false negative for CTI and IM algorithm.

TABLE 3. First level of conformance check results for inductive miner & cti algorithm.

Algorithm	CTI Algorithm		Inductive Miner Algorithm	
Conformance	True Positive	False Negative	True Positive	False Negative
True Positive	17	12	23	28
False Negative	3	7	5	9

CTI algorithm. Results of the second-level conformance check also show the trend that the CTI algorithm is better and efficient as compared to the in-use Inductive Miner algorithm. Figure 7 shows the second level conformance comparison in terms of true positive and false negative for CTI and IM Algorithm.

From the results, when repeating dataset event logs and experiments with the CTI algorithm, better results are dis-

TABLE 4. Second level of conformance check results for inductive miner and CTI algorithm.

Algorithm	CTI Algorithm		Inductive Miner Algorithm (IM)	
Conformance	True Positive	False Negative	True Positive	False Negative
True Positive	7	4	9	8
False Negative	0	9	3	12

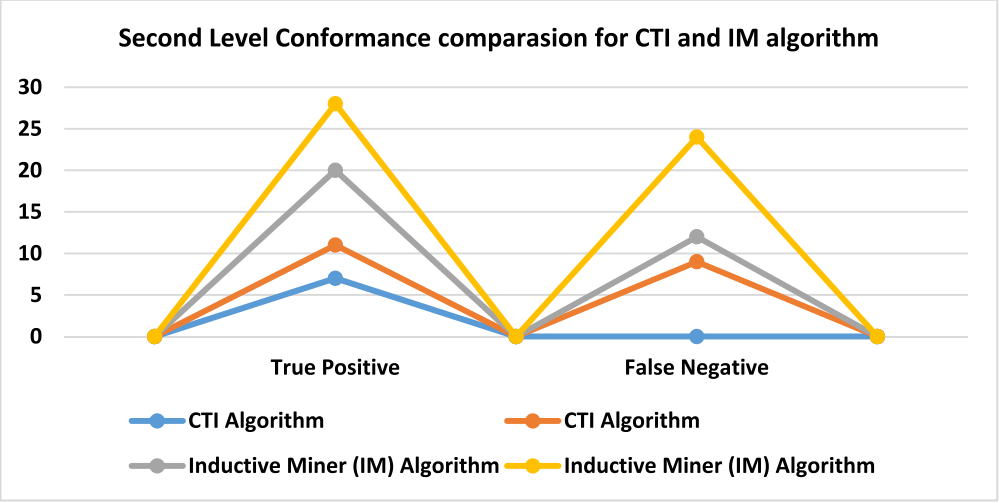

FIGURE 7. Second level conformance comparison in terms of true positive and false negative for CTI and IM algorithm.

played as compared to an existing, in-use algorithm. This validates the proposed method for discovering hidden tasks in ICS system networks. On comparing the process models, the new process shows significantly reduced variances as compared to others.

VI. CONCLUSION

Industrial Control Systems have migrated from being dedicated, air-gapped, centralized infrastructures and have adopted the distributed, corporate systems accessible via the Internet. Although the efficiency, speed, precision quality is increased, this has exposed ICS to the unsecured Internet. This brings the infrastructure open to cybersecurity attacks. By performing process mining for processes and tasks in ICS, tasks and log event discovery were evaluated to determine process behavior attacks and modeling. Conformance checking activities are performed to validate deviations. The demonstrated result shows that the proposed new 'Capturing the Invisible' (CTI) algorithm detected cybersecurity attacks efficiently as compared to other algorithms for Industrial Control Systems. The future work will comprise of prototypes for improving the attack vector classification process by increasing the process mining spectrum for behavior detection and including more, sophisticated, and advanced persistent cyberattacks.

REFERENCES

[1] Y. Hou, J. Such, and A. Rashid, "Understanding security requirements for industrial control system supply chains," in *Proc. IEEE/ACM 5th Int. Workshop Softw. Eng. Smart Cyber-Phys. Syst. (SEsCPS)*, Montreal, QC, Canada, May 2019, pp. 50–53, doi: 10.1109/Sescps.2019.00016.

[2] W. Xu, Y. Tao, and X. Guan, "The landscape of industrial control systems (ICS) devices on the Internet," in *Proc. Int. Conf. Cyber Situational Awareness, Data Anal. Assessment (Cyber SA)*, Glasgow, U.K., Jun. 2018, pp. 1–8, doi: 10.1109/Cybersa.2018.8551422.

[3] G. Guo, J. Zhuge, M. Yang, G. Zhou, and Y. Wu, "A survey of industrial control system devices on the Internet," in *Proc. Int. Conf. Internet Things, Embedded Syst. Commun. (IINTEC)*, Hamammet, Tunisia, Dec. 2018, pp. 197–202, doi: 10.1109/IINTEC.2018.8695276.

[4] K. Hasan, S. Shetty, S. Ullah, A. Hassanzadeh, and E. Hadar, "Towards optimal cyber defense remediation in energy delivery systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–7, doi: 10.1109/GLOBECOM38437.2019.9013517.

[5] G. Williamson. (Jul. 7, 2015). *OT, ICS, SCADA—What's the Difference?* Accessed: Apr. 24, 2020. [Online]. Available: https://www.kuppingercole.com/blog/williamson/ot-ics-scada-whats-the-difference

[6] (Apr. 24, 2020). *Human-Machine Interface in Industrial Networks: Definition, Components, Functions & Examples*. [Online]. Available: https://study.com/academy/lesson/human-machine-interface-hmi-in-industrial-networks-definition-components-functions-examples.html

[7] (Apr. 24, 2020). *OT, ICS, SCADA & PLC*. [Online]. Available: https://www.holmsecurity.com/ot-ics-scada-plc

[8] Waterfall Team. (Dec. 19, 2019). *Infographic: Top 2019 Attacks on ICS*. Accessed: Apr. 24, 2020. [Online]. Available: https://waterfall-security.com/top-2019-attacks-on-ics

[9] R. Zeidler. (Feb. 20, 2020). *What the Explosive Growth in ICS-Infrastructure Targeting Means for Security Leaders*. Accessed: Apr. 24, 2020. [Online]. Available: https://securityintelligence.com/posts/what-the-explosive-growth-in-ics-infrastructure-targeting-means-for-security-leaders

[10] Process Gold. (Mar. 18, 2020). *What is Process Mining: Nearly Everything You Need to Know*. Accessed: Apr. 24, 2020. [Online]. Available: https://processgold.com/process-mining-everything-you-need-to-know

[11] J. Bicknell. (Apr. 24, 2020). *Process Mining Technologies*. [Online]. Available: https://pubsonline.informs.org/do/10.1287/orms.2019.05.01/full

[12] F. Wang, W. Qi, and T. Qian, "A dynamic cybersecurity protection method based on software-defined networking for industrial control systems," in *Proc. Chin. Automat. Congr. (CAC)*, Hangzhou, China, Nov. 2019, pp. 1831–1834, doi: 10.1109/CAC48633.2019.8996244.

[13] M. Soufian, "Towards self-defending control systems in cybersecurity analysis and measures in industrial automation systems," in *Proc. IEEE 26th Int. Symp. Ind. Electron. (ISIE)*, Edinburgh, U.K., Jun. 2017, pp. 1887–1892, doi: 10.1109/ISIE.2017.8001538.

[14] Q. Zhang, C. Zhou, Y.-C. Tian, N. Xiong, Y. Qin, and B. Hu, "A fuzzy probability Bayesian network approach for dynamic cybersecurity risk assessment in industrial control systems," *IEEE Trans. Ind. Informat.*, vol. 14, no. 6, pp. 2497–2506, Jun. 2018, doi: 10.1109/TII.2017.2768998.

[15] X. Lou, K. Waedt, T. Schurmann, H. Kaufhold, V. Watson, and D. Gupta, "Cybersecurity analysis of industrial control system functionality," in *Proc. IEEE Int. Conf. Ind. Cyber Phys. Syst. (ICPS)*, Taipei, Taiwan, May 2019, pp. 73–80, doi: 10.1109/ICPHYS.2019.8780260.

[16] A. L. P. Gómez, L. F. Maimo, A. H. Celdran, F. J. G. Clemente, C. C. Sarmiento, C. J. Del Canto Masa, and R. M. Nistal, "On the generation of anomaly detection datasets in industrial control systems," *IEEE Access*, vol. 7, pp. 177460–177473, 2019, doi: 10.1109/ACCESS.2019.2958284.

[17] T. Abdelghani, "Industrial control systems (ICS) security in power transmission network," in *Proc. Algerian Large Electr. Netw. Conf. (CAGRE)*, Alreris, Algeria, Feb. 2019, pp. 1–4, doi: 10.1109/CAGRE.2019.8713289.

[18] X. Lou, K. Waedt, Y. Gao, I. B. Zid, and V. Watson, "Combining artificial intelligence planning advantages to assist preliminary formal analysis on industrial control system cybersecurity vulnerabilities," in *Proc. 10th Int. Conf. Electron., Comput. Artif. Intell. (ECAI)*, Lasi, Romania, Jun. 2018, pp. 1–8, doi: 10.1109/ECAI.2018.8678949.

[19] X. Li, C. Zhou, Y.-C. Tian, and Y. Qin, "A dynamic decision-making approach for intrusion response in industrial control systems," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 2544–2554, May 2019, doi: 10.1109/TII.2018.2866445.

[20] C. Escudero, F. Sicard, and E. Zamai, "Process-aware model based IDSs for industrial control systems cybersecurity: Approaches, limits and further research," in *Proc. IEEE 23rd Int. Conf. Emerg. Technol. Factory Automat. (ETFA)*, Turin, Italy, Sep. 2018, pp. 605–612, doi: 10.1109/ETFA.2018.8502585.

[21] M. G. Angle, S. Madnick, J. L. Kirtley, and S. Khan, "Identifying and anticipating cyberattacks that could cause physical damage to industrial control systems," *IEEE Power Energy Technol. Syst. J.*, vol. 6, no. 4, pp. 172–182, Dec. 2019, doi: 10.1109/JPETS.2019.2923970.

[22] X. Li, C. Zhou, Y.-C. Tian, N. Xiong, and Y. Qin, "Asset-based dynamic impact assessment of cyberattacks for risk analysis in industrial control systems," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 608–618, Feb. 2018, doi: 10.1109/TII.2017.2740571.

[23] A. Ogundokun, P. Zavarsky, and B. Swar, "Cybersecurity assurance control baselining for smart grid communication systems," in *Proc. 14th IEEE Int. Workshop Factory Commun. Syst. (WFCS)*, Imperia, Italy, Jun. 2018, pp. 1–6, doi: 10.1109/WFCS.2018.8402378.

[24] B. Leander, A. Causević, and H. Hansson, "Cybersecurity challenges in large industrial IoT system," in *Proc. 24th IEEE Int. Conf. Emerg. Technol. Factory Automat. (ETFA)*, Zaragoza, Spain, Sep. 2019, pp. 1035–1042, doi: 10.1109/ETFA.2019.8869162.

[25] M. A. Azad, M. Alazab, F. Riaz, J. Arshad, and T. Abullah, "Socioscope: I know who you are, a robo, human caller or service number," *Future Gener. Comput. Syst.*, vol. 105, pp. 297–307, Apr. 2020, doi: 10.1016/j.future.2019.11.007.

[26] M. Alazab, M. Alazab, A. Shalaginov, A. Mesleh, and A. Awajan, "Intelligent mobile malware detection using permission requests and API calls," *Future Gener. Comput. Syst.*, vol. 107, pp. 509–521, Jun. 2020, doi: 10.1016/j.future.2020.02.002.

[27] M. Alazab, "Profiling and classifying the behavior of malicious codes," *J. Syst. Softw.*, vol. 100, pp. 91–102, Feb. 2015, doi: 10.1016/j.jss.2014.10.031.

AKASHDEEP BHARDWAJ received the Engineering degree in computer science, the master's degree in management, and the Ph.D. degree in computer science. He worked as a Technology Leader and the Head for various multinational organizations. He is currently a Professor of cyber security and digital forensics with the University of Petroleum and Energy Studies (UPES), Dehradun, India. He was an eminent IT Industry Expert with over 25 years of experience in areas, such as cybersecurity, digital forensics, and IT management operations. He also mentors graduate, master's, and Ph.D. students and leads several projects. He has published several research articles. He has coauthored several books. He is certified in cybersecurity, compliance audits, information security, Microsoft, Cisco, and VMware technologies.

FADI AL-TURJMAN (Member, IEEE) received the Ph.D. degree in computer science from Queen's University, Kingston, ON, Canada, in 2011. He is currently a Full Professor and a Research Center Director with Near East University, Nicosia, Cyprus. He is a leading authority in the areas of smart/intelligent, wireless, and mobile networks' architectures, protocols, deployments, and performance evaluation. His publication history spans over 250 publications in journals, conferences, patents, books, and book chapters, in addition to numerous keynotes and plenary talks at flagship venues. He has authored and edited more than 25 books about cognition, security, and wireless sensor networks' deployments in smart environments, published by Taylor and Francis, Elsevier, and Springer. He has received several recognitions and best papers' awards at top international conferences. From 2015 to 2018, he received the prestigious *Best Research Paper Award* from Elsevier Computer Communications Journal and the *Top Researcher Award* for 2018 at Antalya Bilim University, Turkey. He has led a number of international symposia and workshops in flagship communication society conferences. He serves as an Associate Editor and a Lead Guest/Associate Editor for several well-reputed journals, including the IEEE Communications Surveys and Tutorials (IF 22.9) and Sustainable Cities and Society (IF 4.7) (Elsevier).

MANOJ KUMAR received the B.Tech. degree in computer science from Kurukshetra University, the M.Sc. degree in information security and forensics from ITB, Dublin, the M.Tech. degree from ITM University, and the Ph.D. degree in computer science from The Northcap University, Gurugram. He was associated as an Assistant Professor with the Computer Science Department, Amity University Uttar Pradesh, Noida. He is currently working as an Assistant Professor (SG), (SoCS) with the University of Petroleum and Energy Studies, Dehradun. He has over nine years of experience in research and academics. He published over 25 publications in reputed journals and conferences. He is a member of various professional bodies, including the IEEE, ACM, ISTE, and so on. He was a Reviewer for many reputed journals.

THOMPSON STEPHAN received the B.E. degree in computer science and engineering and the M.E. degree in computer science and engineering from Anna University, India, and the Ph.D. degree in computer science and engineering from Pondicherry University, India. He is currently an Assistant Professor with the Department of Computer Science and Engineering, Amity University Uttar Pradesh, Noida, India. His research interests include cognitive radio communications, ad hoc and sensor wireless networks, vehicular communications, and the Internet of Things.

LEONARDO MOSTARDA (Member, IEEE) received the Ph.D. degree from the Computer Science Department, University of L'Aquila, in 2006. He was with the European Space Agency (ESA) on the CUSPIS FP6 project to design and implement novel security protocols and secure geo tags for works of art authentication. He was combining traditional security mechanisms and satellite data. He was Research Associate with the Distributed System and Policy Group, Computing Department, Imperial College London, in 2007. He was working on the UBIVAL EPRC project in cooperation with Cambridge, Oxford, Birmingham, and UCL for building a novel middleware to support the programming of body sensor networks. He was Senior Lecturer with the Distributed Systems and Networking Department, Middlesex University, in 2010. He founded the Senso LAB an innovative research laboratory for building energy efficient wireless sensor networks. He is currently an Associate Professor and the Head of the Computer Science Department, Camerino University, Italy.