Computers & Security 146 (2024) 104070

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

A compliance assessment system for Incident Management process

Alessandro Palma [a,*], Giacomo Acitelli [a], Andrea Marrella [a], Silvia Bonomi [a], Marco Angelini [b]

[a] Department of Computer, Control, and Management Engineering "Antonio Ruberti", Sapienza University of Rome, Via Ariosto 25, Rome, 00185, Italy
[b] Link Campus University, Via del Casale di San Pio V, 44, Rome, 00165, Italy

ARTICLE INFO

Keywords:
Incident management
Security governance
Process compliance assessment
Cost model
Trace alignment

ABSTRACT

The Incident Management (IM) process is one of the core activities for increasing the overall security level of organizations and better responding to cyber attacks. Different security frameworks (such as ITIL and ISO 27035) provide guidelines for designing and properly implementing an effective IM process. Currently, assessing the compliance of the actual process implemented by an organization with such frameworks is a complex task. The assessment is mainly manually performed and requires much effort in the analysis and evaluation. In this paper, we first propose a taxonomy of compliance deviations to classify and prioritize the impacts of non-compliant causes. We combine trace alignment techniques with a new proposed cost model for the analysis of process deviations rather than process traces to prioritize interventions. We put these contributions into use in a system that automatically assesses the IM process compliance with a reference process model (e.g., the one described in the chosen security framework). It supports the auditor with increased awareness of process issues to make more focused decisions and improve the process's effectiveness. We propose a benchmark validation for the model, and we show the system's capability through a usage scenario based on a publicly available dataset of a real IM log. The source code of all components, including the code used for benchmarking, is publicly available as open source on GitHub.

1. Introduction

Security incidents can happen on a daily basis as attackers are developing smarter and can leverage an increasing number of vulnerabilities to intrude and attack their target systems. According to ISO 27035 (ISO/IEC 27035:2013 (E), 2013), a *security incident* is an unwanted or unexpected set of events with a significant probability of compromising business operations and threatening security. Managing incidents is fundamental for every organization to react and promptly contain the consequences of an attack. *Incident Management* (IM) is the process of detecting, reporting, assessing, responding to, and dealing with security incidents.

However, if not properly managed, the IM process may require significant time to coordinate internally with the team and analyze resources (Madigan et al., 2004; Ali et al., 2021). In this scenario, it is crucial for an organization to perform *compliance assessment*, which is the process of evaluating the current state of compliance oversight, management, and related risks in a specific area (e.g., IM) (ISO 19600:2014 (E), 2014). Among the possible compliance analyses, *process compliance assessment* involves comparing the actual implementation of a process with a reference one (e.g., a reference process model from a standard). When the implemented process differs from the reference one, then there exist *process deviations*, which are warnings of unexpected process executions since they make the process non-compliant with the reference one. Such warnings should be controlled as every process entails costs deriving from resources, production, and deployment (Dumas et al., 2013; Palma et al., 2024). It increases significantly in the case of non-compliant processes since fixing a problem highlighted as a warning is more expensive than executing it correctly the first time (Madigan et al., 2004).

Problem Statement. As stated by Siponen and Willison (2009), it is difficult to manage the compliance of processes with standards as (i) it is difficult to measure the compliance degree of a process; (ii) there exist various application scenarios with different constraints (e.g., the healthcare domain is different from the finance domain) and (iii) typically the standards are intentionally general to be widely applicable (e.g., administrative guidelines just elicit what should be done, but not how). Considering these aspects, evaluating the compliance of an IM process with the standards through classic manual approaches raises several issues: subjective bias is possible in each step of the assessment, and different sensibilities from different auditors to similar situations may influence the evaluation. In addition, it can be costly, slow, and cumbersome (Carmona et al., 2018a). Finally, most of the literature

* Corresponding author.
E-mail address: palma@diag.uniroma1.it (A. Palma).

https://doi.org/10.1016/j.cose.2024.104070
Received 7 August 2023; Received in revised form 21 June 2024; Accepted 18 August 2024
Available online 22 August 2024

analyzes the incidents based on their aggregated cost, considering different information such as resources, personnel, and impacts: in these cases, it is difficult to identify the weak parts of the IM process (Shinde and Kulkarni, 2021; Kieninger et al., 2013; Glogovac et al., 2019).

Research Methodology. To mitigate these limitations, we introduce a compliance assessment system to automatically estimate the cost of incidents based on the causes of non-compliance, intended in terms of process deviations. In pursuit of this objective, our methodology begins by formally defining the compliance assessment model designed to quantitatively assess the cost due to non-compliance with security standards. This involves initially establishing a deviation taxonomy classifying the *process deviations* by leveraging trace alignment (Adriansyah et al., 2011; De Leoni and Marrella, 2017) to identify the possible non-compliance issues during the IM process. Then, we introduce a quantitative compliance assessment model defining the non-compliance cost associated with each incident. The next step is the design of a system that supports the auditors in assessing the compliance of the implemented IM process with a reference process model. It comprises two core parts: in the former, we leverage trace alignment to identify the deviations in the implemented IM process. In the latter, we map the proposed compliance assessment model into the identified deviations and automatically evaluate the non-compliance cost of incidents according to three approaches: causal probability, linear regression, and extra-tree regression. The granularity of the assessment at the deviation level helps to be more effective than the state-of-the-art solutions, which analyze compliance at the whole incident level. This supports an auditor in automatically prioritizing the most impactful and costly causes of non-compliance and reducing their analyses to a manageable number. To validate the effectiveness of the proposed solution, we perform a validation using five ground truths from the state-of-the-art. In addition, we present a usage scenario to show the capabilities of the proposed system. The source code of all components, including the code used for benchmarking and usage scenario, is publicly available as open source on GitHub[1]

Contributions. In summary, this paper contributes the following:

- the design of a general deviations taxonomy and a cost model to classify the different issues arising from non-compliance with the reference process model;
- the design and implementation of a compliance assessment system to automatically support the auditor in assessing the causes of non-compliance;
- a benchmark to validate the proposed taxonomy and cost model applied to the IM process;
- a usage scenario to show the system's capabilities for ISO 27035 assessment.

Paper Structure. The paper is organized as follows: Section 2 introduces the fundamental concepts of the IM process and its compliance assessment. The related work is reported in Section 3, while Section 4 presents an illustrative example of the ISO 27035 process. The model is described in Section 5, and the compliance assessment system is detailed in Section 6. Section 7 presents the validation results of the proposed model, while Section 8 shows a usage scenario of the compliance assessment system. Finally, Section 9 summarizes the paper, highlighting limitations and promising research directions, and Section 10 concludes the paper.

2. Background

In this section, we report some foundational notions of the IM process, its assessment, and basic process mining concepts used in the rest of the paper.

[1] https://github.com/Ale96Pa/ComplianceAssessmentSystem

2.1. Incident Management process

The security standards documenting the Incident Management (IM) process (i.e., ISO/IEC 27035:2013 (E), 2013, ITILv4, 2019, ENISA, 2010, and National Institute of Standards and Technology, 2021) report five necessary phases to manage incidents correctly. The preliminary step is *planning and preparation*, in which the organization establishes the policies and the competent teams to deal with incidents. In this phase, the organization decides how to log the incidents and which information to include. The *detection and reporting* phase consists of spotting and logging all the activities that may turn into incidents (e.g., suspicious accesses). Once a potential incident is detected, an auditor performs the *incident assessment* which consists of categorizing the incident based on the impacted IT or business areas and determining the priority based on impact and urgency. Based on this assessment, the incident is consequently routed to a technical operator with the relevant expertise responsible for *incident response*, i.e., providing mitigation actions that the organization must implement to contain the damages. The final step is *incident closure*, in which the organization systematically reviews the risks due to the incident experienced and acknowledges the impacted users about it.

The security standards also describe the workflow of the IM process, that is the order in which these phases must be performed. It may vary among the different standards. For example, while ISO 27035 and ITIL report the incident response to be executed only after the incident assessment, NIST provides an iterative loop between assessment and response, with the results of the two phases informing each other.

2.2. Incident Management process compliance assessment

To put in place a security process, an organization typically refers to a *reference security standard* and implements all the documented guidelines included in it. When referring to a process (as for the IM), the standards describe how the process must be performed, defining a *reference process model* that the organization must comply with. To support the organizations in implementing the security processes, several IT service management systems (ServiceNow, 2023; Solarwind, 2021) exist that automate the ticketing system (Gohil and Kumar, 2019), which consists of associating a ticket for any problem that the organization must solve. In the case of the IM process, a user opens a ticket to report the detected problem(s) potentially leading to an incident. This allows tracking the entire process life cycle from opening to closure. We refer to the information logged through such systems during an IM process as *incident management log*. To be effectively used to assess the process compliance of an organization, the incident management log must include the incident identifiers, the phases performed for each incident, their timestamps, the team that performed each phase, and the actors who detected, opened, resolved, and closed the ticket (Accorsi, 2009). Additional information can be present to enrich further the analysis, such as Service Level Agreement (SLA) violations, incident location, category, priority, and mitigation actions that are put in place.

IM process compliance assessment in practice. When an auditor assesses the compliance of the IM process implemented by an organization, s/he typically analyzes the incident management log according to checklists provided by the reference standard. To do so, it is fundamental that the auditor is an expert on the used reference process. Examples of compliance requirements an auditor must check are ISO/TC 9001 (2014), van der Kleij et al. (2022):

- process definition and documentation;
- accountability of process performance and compliance;
- consistency of organization plans and implementation;
- minimization of redundant non-value-adding activities.

An auditor determines the compliance level of each requirement with the reference standard according to predefined scales (e.g., compliant, partially compliant, and not compliant); then, s/he aggregates the results (e.g., average) for evaluating an organization's overall compliance. On the one hand, manually performed assessment has some advantages as the human auditor can handle complex and fine-grained decision-making tasks. In addition, an organization with an internal process auditor benefits from having someone continually dedicated to monitoring the IM process. On the other hand, some drawbacks derive from manual assessment: it requires more time with respect to automated approaches, and the probability of errors increases, as, for example, it is easy for a human to transpose numbers or make other similar mistakes (e.g., misinterpretation).

2.3. Process mining

Process Mining (Van Der Aalst, 2016) is a family of data analysis techniques that encompasses several sub-disciplines, such as process discovery (Augusto et al., 2018), conformance checking (Carmona et al., 2018a), deviance analysis (Nguyen et al., 2014), and predictive process monitoring (Márquez-Chamorro et al., 2017).

Process mining concentrates on the actual execution of processes, as reflected by the footprint of reality logged by an organization's information systems. The main input is an *event log* (e.g., incident management log), which is analyzed to extract insights and recurrent patterns about how processes are executed. Event logs consist of *traces* (e.g., incidents instances). A trace consists of the sequence of *events* (e.g., incidents phases) logged during the execution of an individual instance of a process. In the context of this paper, we leverage process mining, in particular trace alignment, to identify process deviations and assess their severity in the IM process automatically.

3. Related work

The related work is organized following three main areas that intersect with our contribution: research related to process compliance assessment, its automation and approaches related to cost models for compliance analysis.

Process Compliance Assessment. Process Compliance Assessment is the task of analyzing and evaluating an organization's process implementation. The auditor performing this job typically makes interviews and manual analysis to map evaluation metrics to requirements provided by the reference model (ISO 19600:2014 (E), 2014; Siponen and Willison, 2009; ISO 37301:2021 (E), 2021; González-Granadillo et al., 2021). Manual approaches raise several issues (e.g., auditors' bias) during the evaluation. Indeed, van der Kleij et al. (2022) performed an expert-driven evaluation to determine the decisional tasks of an auditor during the IM process. They found there is a lack of attention to the details, while only macro-activities are checked. For this reason, some works proposed methodologies to support the auditor during the process assessment (Angelini et al., 2020; El Kharbili, 2012; Angelini et al., 2022).

Among the most relevant approaches, Ly et al. (2012) propose a framework to specify general compliance criteria and to check whether the actual process meets them. They leverage ontologies and formal rules to check compliance with given constraint specifications. Similarly, Liu et al. (2007) and Arsac et al. (2011) propose methodologies based on model-checking techniques to formally verify the compliance of the process adopted by an organization with business requirements, and the latter also provides a supporting tool. Kabaale et al. (2018) propose another approach that leverages formal process verification to process compliance assessment. They extract process requirements from the standard documents, translate them into logical axioms and evaluate the compliance of the modeled requirements with an ontology-based approach. Caron et al. (2012) identify the events that might adversely affect the business objectives and estimate the risk severity. They translate business directives into one or more specific controls, which map on generic rule patterns (e.g., segregation of duties, activity existence, and arithmetic derivation patterns). Mouratidis et al. (2023) introduce a conceptual model for the incident response that integrates concepts from different domains, such as security requirements, forensics, threat intelligence, critical infrastructures, and incident handling. Additionally, He et al. (2022) and Naseer et al. (2023) propose two agile methodologies to replace the classic sequential IM process framework. The former adds feedback activities at the early steps of the IM process, and the latter introduces the impact of big data analytics in the response phase. Finally, Alfaadhel et al. (2023) contribute to a compliance assessment system based on checking the coverage of security controls.

Contrary to those approaches, we do not consider constraints coming from fixed rules or security controls as they could be complex to manage manually and not general enough. Instead, we propose leveraging automatic inference of the actual process issues from an event log, using a cost model to evaluate their severity and better assess their impacts.

Automated Process Compliance Assessment. To automate process compliance assessment, a suitable area to leverage is conformance checking (Carmona et al., 2018a), a process mining (Van Der Aalst, 2012) discipline to detect anomalies in business processes and assess the compliance degree with the *fitness* (Adriansyah et al., 2011; De Leoni and Marrella, 2017), which measures the number of process deviations. Indeed, some works use it for auditing and risk assessment purposes to detect anomalies during the process execution (Silalahi et al., 2022; Kothandapani, 2023).

Vanden Broucke et al. (2013) describe a framework's architecture to assess a process model's goodness from a quantitative perspective, leveraging different conformance checking metrics. Although comprehensive, standardizing the analysis of deviations over multiple metrics is a challenging task that they highlight as an open problem. To mitigate this limitation, we propose a cost model to quantify the impact of the deviations causing non-compliance. Varela-Vaca et al. (2011) present a security risk information framework that integrates the asset value for each process activity in terms of confidentiality, integrity, and availability, and estimates the risk value as a linear combination of asset value, activity frequency and threat extent, Although automated, their work requires the human evaluation of the asset values prior to the assessment, while our system uses automatic approaches also to estimate the cost of non-compliance issues considering the process log. De Leoni et al. (2014) propose a technique to check the conformance of data-aware process models. The innovative idea inspiring this work is the consideration of the context of the process in addition to the control-flow perspective. As they state, the work only focuses on the fitness aspect of conformance, while we consider a novel cost model to evaluate the non-compliance issues. Bernardi et al. (2018) provide an assessment methodology using state-of-the-art techniques based on process mining. They leverage trace alignment to calibrate and validate the performance scenario, which is assessed through the notion of fitness, and the costs are simulated based on mean execution time. Waspada et al. (2022) adapt trace alignment for real-time environments by leveraging a graph-based approach. This enables early decision-making about the compliance of the process with a reference process model. Ghanem et al. (2023) propose an automated security compliance framework that re-uses the expertise knowledge in similar scenarios to avoid redundant checks and improve the execution time of compliance assessment.

Differently from the presented approaches, we propose a solution that not only evaluates the degree of compliance with the reference process model, but it supports a more informed process assessment by prioritizing the non-compliance issues and assigning them an appropriate cost, representing their impact on the compliance.

Cost Models for Process Assessment. Although several works addressed the problem of process compliance assessment from different

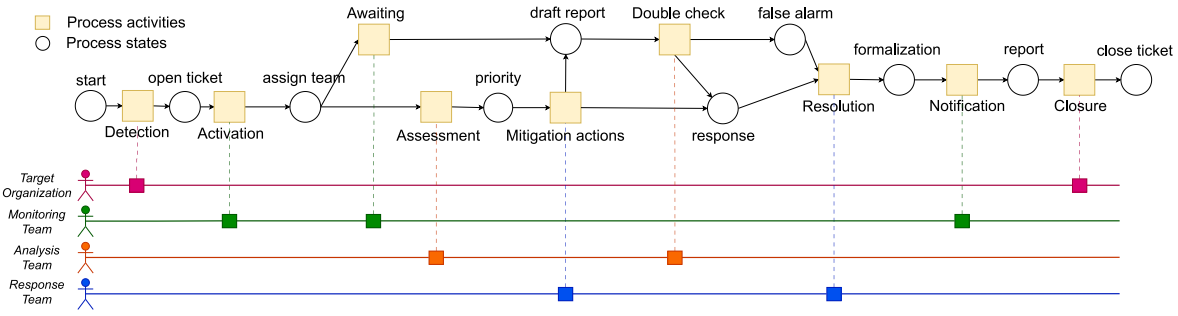


Fig. 1. Petri Net of the Incident Management process model from ISO 27035:2013. It starts with the detection of an incident from the organization and reported to the monitoring team which proceeds with the incident activation (i.e., the opening of a ticket). Then, the incident is either marked as a false alarm or must be assessed twice by the analysis team. Finally, the response team resolves the incident, and the monitoring team notifies the organization, responsible for closing it.

perspectives, few coped with the formalization of cost models to prioritize the mitigation of non-compliance issues. Moura et al. (2006) and Kieninger et al. (2013) estimate the non-compliance cost considering the traces as a whole. The former models the cost of service level agreement violations, in which the cost is estimated as a function of the duration of each SLA violation. The latter quantifies the negative impact of incidents through a simulation-based approach: they identify the business cost components and the process-related metrics and estimate the incident costs by summing up the costs of the different business components. Contrary to these works, we associate the costs with individual deviations so that the auditor has more control over the mitigation of non-compliance.

Pramanik et al. (2005) propose a general deviation-based strategy for the synthesis of tolerances. As they state, the exact values for the features of any part of the model have been an ad-hoc process, mostly experience-based. Contrary, our model estimates the deviation parameters through automated approaches, which avoid their manual assignment to each deviation. Pascual et al. (2009) and Sarkar and Saren (2016) develop decision support tools to inspect errors and warranty costs. They evaluate the cost of the different parts of a system (e.g., electric components, control system, motors) considering failures, labor cost, production cost, and shutdown actions. These costs are then linearly summed to prioritize the most critical system components. Glogovac et al. (2019) leverage the PAF model (i.e., the cost associated with Preventive, Appraisal, and Failure) to design a cost model to prioritize process improvement opportunities. It is based on mathematical models in which the main idea is to weigh the previously categorized costs with the number of non-compliance causes within a given period. Differently from these solutions, in our cost model, we do not assign a cost to the activities nor to system components, but we go deeper into the analysis by considering the type of process deviations.

Beyond these works, other approaches exist that are related to specific domains, such as software development (Jadhav et al., 2022), injection molding (Kazmer et al., 2023), healthcare (Vanounou et al., 2007; Santos et al., 2023), that are difficult to generalize to our problem as they use features specific of the application domain. To the best of the authors' knowledge, this work is the first addressing the problem in the IM context as the literature mainly focuses on trace clustering (De Weerdt et al., 2012; Bertrand et al., 2023), security performance (Wibawa and Ramantoko, 2022) and financial damage (Accorsi and Stocker, 2012; Aldasoro et al., 2022). However, they do not perform IM compliance assessment, nor do they detect, evaluate, or analyze the process errors, as our approach does.

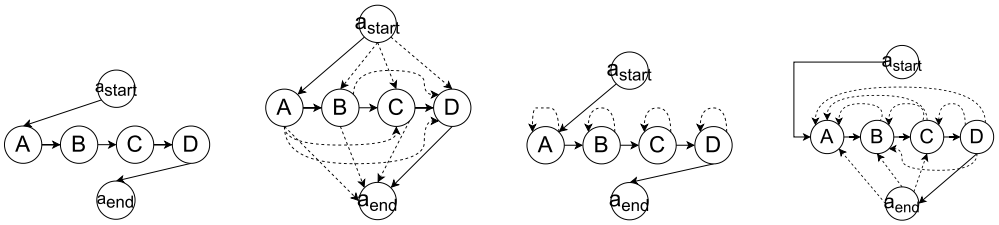
4. Illustrative example: ISO 27035

In this section, we show an example of the IM process from ISO 27035:2013 to describe the problems an auditor may encounter during the IM process compliance assessment. The ISO 27035:2013 IM process is reported in Fig. 1. It is represented as a Petri Net (Petri, 1966), i.e., a directed graph composed of two types of nodes: places (represented as circles) which are states of the process, and transitions (represented as squares) which are activities of the process. The starting activity is incident detection, i.e., a user from the organization who observes suspicious activities and opens a ticket to alert the monitoring team. Once a ticket is opened, the organization is in a state of emergency, and the incident is activated. At this point, either the organization needs a third-party company to address the problem, so it awaits further analyses monitored by the monitoring team, or an appropriate internal analysis team is selected to assess the problem and determine the incident priority. In the former case, once the third-party company provides details about assessment and mitigation actions to put in place, the internal analysis team should double-check such information and either respond to the incident or mark it as a false alarm. In the latter case, once the internal team assesses the incident, the response team puts in place mitigation actions that should be double-checked to mark the incident as resolved or as a false alarm. In case of an extreme emergency, the double-checking step could be skipped, and the team immediately responds to the incident. After the incident response, its resolution is formalized into an appropriate document, the team notifies all the impacted users about the recovery of the normal state, and the incident (and corresponding ticket) is closed.

A real implementation of the IM process is reported in a publicly available dataset containing data from the audit system of the ServiceNow™ (ServiceNow, 2023) platform used by an IT company (Amaral et al., 2019). The log is anonymized for privacy issues and contains 24918 incidents and, for each incident, descriptive features related to the IM process (i.e., for each incident, its identifier, the different phases it is composed of, and the timestamp and the identifier of the people in charge of each phase), incident classification (i.e., incidents priority, category, and location), and incident diagnosis (i.e., impacted Service Level Agreements and the number of times the caller rejected the resolution).

In this example, if an auditor wants to assess the compliance of the implemented IM process manually, then s/he should consider an ISO-based checklist (ISO/TC 9001, 2014) and investigate the log that, as in this case, can be very large. Although different log analysis tools exist (Vaarandi, 2005), they do not correlate the log data to the reference process model (i.e., Fig. 1). Thus the auditor must manually extract the information s/he needs to evaluate, for example, if a different execution of the process is allowed or not. This task can be very cumbersome, time-consuming, and errors prone because the following tasks are all in charge of the auditor:

- Correct interpretation of the IM log attributes: if some attribute is misunderstood, this may result in an inaccurate estimation of the compliance level (e.g., the provided log describes that the incident *urgency* is related to *impact* and *priority*, without specifying how).



(a) Simple reference process model. (b) Skipping deviations. (c) Repetition deviations. (d) Mismatch deviations.

Fig. 2. (a) A simple example of a process model and its related process deviations: (b) skipping, (c) repetition, and (d) mismatch.

- Information inference from the IM log: it may happen that some incidents are not documented (e.g., privacy, missing documentation), thus the auditor should evaluate if process deviation may be tolerated for the compliance assessment or not. Without enough data, the auditor should relate how that deviation impacts the whole process (i.e., the whole log). If not supported by automatic systems, this task is time-consuming and hard to perform.
- Evaluation of the impact of non-compliance with the reference model: typically, incidents have a cost associated with their impact, damages, and resources necessary for responding. Defining how much of this cost is caused by non-compliance with the reference model can result in a too coarse-grained estimation if the correlation is inaccurate, data is missing, or the chosen cost model does not consider the IM process execution.

In the rest of this paper, we propose a compliance assessment system to support the auditor during the process compliance assessment, automatically relate the log information to the reference process model, and finally quantify the severity of each deviation from the reference process model.

5. Compliance assessment model

In this section, we define the compliance assessment model and we introduce the proposed deviation taxonomy and its related cost model. We consider an organization O that wants to assess its internal processes. Each process $\mathcal{P}$ comprises a set of activities that must be executed in a specific (potentially partial) order. Thus, we model $\mathcal{P} = (A, E)$ as a directed graph, where A is the set of all activities in the process plus the two fictional activities a_{start} and a_{end} which represent the beginning and end of the process, and E represents the set of precedence relationships between pairs of activities, i.e., an edge $e_{i,j} \in E$ means that activity a_i is executed immediately before a_j.

Every process $\mathcal{P}$ can have multiple *instances*, and we denote the generic ith instance of $\mathcal{P}$ as $\mathcal{P}^i$. Given an instance $\mathcal{P}^i$ of a process $\mathcal{P}$ (actually implemented by O) and a *reference process* $\mathcal{P}_{ref}$ (e.g., defined by an external third party or by a standard)[2], we say that $\mathcal{P}^i$ *is compliant with* $\mathcal{P}_{ref}$ if $\mathcal{P}^i$ includes the same set of activities considered by $\mathcal{P}_{ref}$ and activities are executed following the same order. By extension, we say that a process $\mathcal{P}$ *is compliant with* $\mathcal{P}_{ref}$ if any instance $\mathcal{P}^i$ of $\mathcal{P}$ is compliant with $\mathcal{P}_{ref}$.

Definition 1. Let $\mathcal{P}^i = (A^i, E^i)$ and $\mathcal{P}_{ref} = (A_r, E_r)$ be respectively an instance of a process implemented in O and its current reference process. We say that $\mathcal{P}^i$ *deviates from* $\mathcal{P}_{ref}$ if: (i) they are not executing the same set of activities (i.e., $A^i \neq A_r$) or (ii) there exists at least a pair of activities that is not executed in the correct order (i.e., $E^i \neq E_r$) or (iii) there exists at least an activity a_i that is executed multiple times.

[2] Without loss of generality, we assume that the reference process $\mathcal{P}_{ref}$ does not contain self-loop. This assumption can be easily removed by considering any previously mentioned graph as a multi-graph.

By extension, we say that a process $\mathcal{P}$ *deviates from* $\mathcal{P}_{ref}$ if there exists at least one instance $\mathcal{P}^i$ of $\mathcal{P}$ that deviates from $\mathcal{P}_{ref}$. Let us note that if $\mathcal{P}$ deviates from $\mathcal{P}_{ref}$ it means that there is at least one difference in the two graphs representing the processes, i.e., there is at least one *deviation*. Our goal is to quantify the deviations of $\mathcal{P}$ from $\mathcal{P}_{ref}$ and measure their impact in terms of non-compliance cost. To this aim, we denote as $C_{NC}(\mathcal{P}^i)$ the *non-compliance cost* of a specific instance $\mathcal{P}^i$ of $\mathcal{P}$ and $C_{NC}(\mathcal{P})$ the non-compliance cost of the overall process, where $C_{NC}(\mathcal{P}) = f(C_{NC}(\mathcal{P}^1), C_{NC}(\mathcal{P}^2), \dots C_{NC}(\mathcal{P}^k))$.

Any instance $\mathcal{P}^i$ is constructed by *logging* all activities executed during the ith instance of $\mathcal{P}$ in the *activity log* L. We assume that L is a (potentially infinite) execution of $\mathcal{P}$ composed of a sequence of tuples $t_1, t_2, \dots$ where each entry must contain at least the following information (i) the identifier i of the current instance of $\mathcal{P}$, (ii) a *timestamp* ts and (iii) the activity a that has been executed. Additional information may also be included in the log to characterize better the activity execution (i.e., $t_i = \langle id, ts, a, \dots \rangle$).

We take the *log consistency* assumption in which activities are logged consistently with the reference process, i.e., the activity names in the log are the same as the reference process ones. Thus, there does not exist any activity executed by $\mathcal{P}$ and logged in L that is not in A_r, and we do not consider deviations due to the execution of unknown activities.

5.1. Deviation taxonomy

The first contribution of the compliance assessment model is the definition of a deviation taxonomy to classify the different non-compliance issues. For the sake of explanation, let us consider the (simple) reference process $\mathcal{P}_{ref}$ represented in Fig. 2(a). It is composed of four activities (plus the fictional one for starting and ending the process) that need to be executed in sequence (i.e., $\mathcal{P}_{ref} = (A_r, E_r)$ where $A_r = \{A, B, C, D, a_{start}, a_{end}\}$ and $E_r = \{e_{start,A}, e_{A,B}, e_{B,C}, e_{C,D}, e_{D,end}\}$.

Given the *log consistency* assumption, a deviation may occur only when there exists at least one instance of $\mathcal{P}$ where the set of edges E^i characterizing the actual instance is different from the set of edges considered by the reference process E_r. Thus, in our case, the set of deviations $D_{(\mathcal{P}^i, \mathcal{P}_{ref})}$ is defined as the set of edges that appears in E^i but does not appear in E_r and vice-versa (i.e., $D_{(\mathcal{P}^i, \mathcal{P}_{ref})} = E^i \veebar E_r$).

By extension, the set of deviations of the whole process $\mathcal{P}$ is the union of the deviation of every instance of $\mathcal{P}$ (i.e., $D_{(\mathcal{P}, \mathcal{P}_{ref})} = \bigcup_i D_{(\mathcal{P}^i, \mathcal{P}_{ref})}$). Let us observe that the same deviation may occur multiple times in the same process (and even in the same instance), and we want to preserve this information. Thus, D is a multi-set rather than a simple set with no duplicates. In the following, when not ambiguous, we simplify the notation by writing D instead of $D_{(\mathcal{P}, \mathcal{P}_{ref})}$.

After these considerations, we can introduce our first contribution: a general taxonomy of process deviations. Given a process instance $\mathcal{P}^i$ and the reference process $\mathcal{P}_{ref}$, let $path(a_j, a_k)$ in a process $\mathcal{P}$ be a sequence $a_j, a_{j+1}, a_{j+2}, \dots a_k$ such that for any pair of adjacent activities there exists an edge in $\mathcal{P}$. Then, any deviation in D must be in one of the following categories:

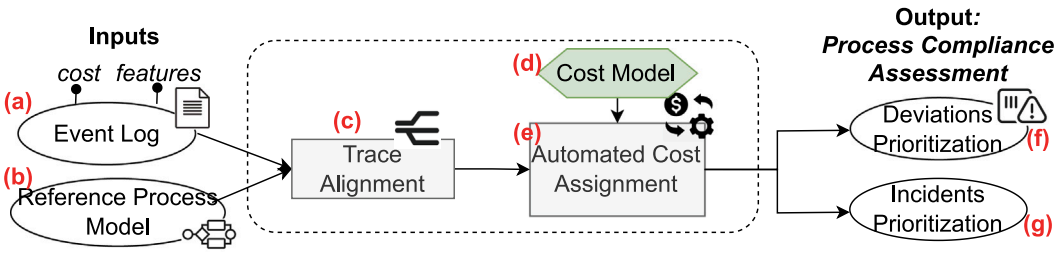


Fig. 3. Compliance Assessment System architecture. The inputs are (a) an event log including incidents cost and features (e.g., personnel involved) and (b) a reference process model (e.g., modeled as a Petri net). They are processed by a trace alignment module (c) which identifies process deviations and assigns a fitness value to each incident. Once deviations are computed, the cost model (d) allows to evaluate the non-compliance cost of incidents based on the cost of each deviation which is assigned automatically (e). Based on the non-compliance cost, deviations (f) and incidents (g) are prioritized based on their severity.

1. *skipping deviations* $D_{skip} \subseteq D$ (dotted edges in Fig. 2(b)): this set represents edges belonging to the actual process instance, not included in the reference process and that would require *to skip/delete* an activity in the reference model. More formally $D_{skip} = \{e_{j,k} \in E^i \setminus E_r \mid \exists path(a_j, a_k) \in \mathcal{P}_{ref}\}$. As an example, the edge $e_{A,C}$ in Fig. 2(b) is a skipping deviation as it does not exist in E_r, in which we have the path $\{A, B, C\}$ (see Fig. 2(a)), i.e., moving from activity A to C involves skipping activity B.
2. *repetition deviations* $D_{rep} \subseteq D$ (dotted edges in Fig. 2(c)): this set represents loops in the execution of the same activity not foreseen by the reference process. More formally, $D_{rep} = \{e_{i,i} \in E^i \mid e_{i,i} \notin E_r\}$.
3. *mismatch deviations* $D_{mis} \subseteq D$ (dotted edges in Fig. 2(d)): this set represents the edges that indicate the execution of a given set of activities in a different order with respect to what is established from the reference process. More formally, $D_{mis} = \{e_{j,k} \in E^i \setminus E_r \mid \exists path(a_k, a_j) \in \mathcal{P}_{ref} \wedge \nexists path(a_j, a_k) \in \mathcal{P}_{ref}\}$. For example, $e_{B,A}$ in Fig. 2(d) is a mismatch deviation as in the reference process A should be executed before B.

Let us note that this taxonomy is exhaustive under the *log consistency* assumption as it allows classifying any deviation in one of the three defined categories. As an example, let us consider an activity log $L = \{\langle i, 5, B\rangle, \langle i, 7, B\rangle, \langle i, 8, D\rangle, \langle i, 10, C\rangle\}$ and the simple reference process model of Fig. 2(a). Analyzing L it is possible to derive the instance $\mathcal{P}^i = (A_{ref}, E^i)$, where $E^i = \{e_{start,B}, e_{B,B}, e_{B,D}, e_{D,C}, e_{C,end}\}$ and we have the following observed deviations: $D^i_{skip} = \{e_{start,B}, e_{B,D}, e_{C,end}\}$ $D^i_{rep} = \{e_{B,B}\}$ and $D^i_{mis} = \{e_{D,C}\}$ with $D^i = D^i_{skip} \cup D^i_{rep}, \cup D^i_{mis}$.

5.2. Cost model

Knowing only the process deviations could not be enough to prioritize interventions on them. One of the main aspects related to the incidents is their cost. A cost component that is relevant for the analysis of the IM process is the one due to the consequences of an incident, as it is an indicator of the damages caused to an organization (Romanosky, 2016). Very often, this cost may be caused by non-compliance with security standards (Madigan et al., 2004): in that case, we refer to such a cost component as *non-compliance cost*, and a process auditor may be interested in analyzing how this cost is distributed among the different non-compliance causes (i.e., process deviations). This analysis requires a significant effort to map the non-compliance cost to every non-compliance issue. To this aim, we define a cost model to assign a cost value to any deviation observed in the actual (instance of the) process.

Definition 2. Let $\mathcal{P}^i = (A^i, E^i)$ be an instance of the actual process $\mathcal{P}$ and let D^i be the set of deviations observed during the current instance. We define the *cost function* $c(e_{i,j})$ as the function that assigns a numerical cost to any deviation, 0 otherwise.

Then, we define the *non-compliance cost* of a given instance $\mathcal{P}^i$ as:

$$C_{NC}(\mathcal{P}^i) = \sum_{e_{j,k} \in D^i_{skip}} c(e_{j,k}) + \sum_{e_{j,j} \in D^i_{rep}} c(e_{j,j}) \frac{\text{count}(e_{j,j}, D^i_{rep})}{|E^i|} + \sum_{e_{j,k} \in D^i_{mis}} c(e_{j,k}) \frac{\text{count}(e_{j,k}, D^i_{mis})}{|E^i|}, \quad (1)$$

where $\text{count}(e_{i,i}, D_x)$ is the number of edges $e_{i,i} \in D_x$. The cost function $c(e_{i,j})$ weights the impact of the different deviations on the non-compliance cost of the whole incident. Let us note that skipping deviation may appear just once because an activity is either skipped or not. In contrast, repetition and mismatch deviations may be repeated more than once in a process instance, and therefore, they need to be counted to evaluate their impact on the overall instance. In the proposed compliance assessment system, we automatically infer this function through regression and probabilistic approaches (see Section 6.3), although an auditor may choose to manually define such a function or fine-tune the automatic assignment the system proposes. The rationale of Eq. (1) is that it accumulates the *incidence* of each skipping, repeated, and mismatched activity, which is more fine-grained than existing cost functions which typically assign the cost based on global trace features (e.g., duration), disregarding its structure.

Let us consider again the example discussed in the previous paragraph (Fig. 2(a) and log L) and the following cost function:

$$c(e_{i,j}) = \begin{cases} 0.5 & e_{i,j} \in D^i_{skip} \\ 1 & e_{i,j} \in D^i_{rep} \\ 3 & e_{i,j} \in D^i_{mis} \end{cases}$$

Then, the cost associated to $\mathcal{P}^i$, given the cost function above is:

$$C_{NC}(\mathcal{P}^i) = 0.5 + 0.5 + 0.5 + (1 \times \frac{1}{5}) + (3 \times \frac{1}{5}) = 2.3$$

6. Compliance assessment system

To make the compliance assessment analysis operative, in this section, we present our process compliance assessment system to identify and prioritize deviations between an actual implemented process $\mathcal{P}$ and a reference process $\mathcal{P}_{ref}$. The system automatically identifies process deviations and supports the process auditor in the definition and estimation of the cost function $c(e_{i,j})$ by automatically proposing specific costs to any identified deviation based on the set of additional information available in the incident management log. The system architecture is reported in Fig. 3.

6.1. Input and output

The system takes as input a representation of the reference process $\mathcal{P}_{ref}$ (in particular, our tool accepts as input a Petri Net (Petri, 1966)). It is the representation of the IM workflow as provided by security standards (e.g., ISO/IEC 27035:2013 (E), 2013), typically composed of

states, describing a specific state of the process, and *activities*, specifying the types of action to move among different states. In this way, the reference process model describes all the possible workflows that the organization can perform being compliant with the reference security standard. The other input is the incident management log L enriched with the non-compliance cost $C_{NC}(\mathcal{P}^i)$ for any given process instance $\mathcal{P}^i$. Let us remind that a process instance defines a trace, which is the sequence of events logged during an individual execution of a process. It must contain the incident IDs, activities for each incident, and their timestamp as minimum requirements. Moreover, it may include potentially additional features (e.g., personnel IDs, resources involved, incident priority). In the most general case, the non-compliance cost is calculated, leveraging state-of-the-art solutions (Kieninger et al., 2013; Dumas et al., 2013), as the person-hours of *additional* work and can be easily retrieved by the number of people and the duration of extra activities during the incidents. In some cases, it can be estimated simply as the *duration* of such extra activities, as it is a key performance indicator for business processes (Van Der Aalst, 2013). In other cases, more sophisticated non-compliance costs may depend on the log features (e.g., in an IT environment, the downtime service can be considered). Thus, we assume that a rough estimation of the non-compliance cost is always possible in any valid log, as the minimum requirement in the log is the timestamp of each log entry, commonly present. Let us note that the required input is usually available to the organization. Indeed, the incident management log can be collected manually or it can use automatic platforms for digital workflows (e.g., ServiceNow, 2023), while the reference process can be easily defined using appropriate notations (e.g., BPMN (White, 2004), WorkFlow net (Salimifard and Wright, 2001)) by the process auditor (Sonteya and Seymour, 2012).

Concerning the output, the system produces a statistical report. It includes the following information: (i) *Distribution of the non-compliance cost among the traces*: it supports the auditor in the analysis of the evolution of the non-compliance issues among different traces. This indicates the overall compliance level of the IM process under analysis. (ii) *Prioritization of the deviations based on their costs*: the non-compliance cost associated with each deviation supports the auditor in understanding their severity and prioritizing mitigation actions accordingly. (iii) *Prioritization of traces that cause the most significant issues*: the non-compliance cost associated with each incident supports the auditor in detecting the most critical incidents. In this way, s/he can perform a deeper investigation only for a manageable number of selected incidents.

6.2. Trace alignment for deviations identification

The first module of the system is the trace alignment component, which identifies deviations that cause non-compliance. An alignment between a log trace and a process model is a pairwise matching between activities recorded in the log and activities allowed by the model. Trace alignment automatically detects deviations in each trace and classifies them. The trace alignment module can also provide a quantification of a trace level of compliance by computing the state-of-the-art metric *fitness* (Adriansyah et al., 2011; De Leoni and Marrella, 2017). The fitness is expressed with a value between 0 (no log activity matches with the reference model) and 1 (all log activities match the model) and measures the overall number of deviations, without considering different costs for different types of deviation. The trace alignment module automatically produces as output the set of deviations D^i for each process instance $\mathcal{P}^i$, along with the corresponding fitness values. In our implementation, the trace alignment component is implemented by leveraging the pm4py library (Berti et al., 2019).

6.3. Automated costs assignment

The goal of this module is the automatic assignment of the individual non-compliance costs to each deviation. We propose one probability-based (exploiting two different settings) and two regression-based approaches (namely linear regression and tree regression).

Probabilistic approach. This approach leverages causal probability (Skyrms, 1982) as it characterizes the relationship between cause and effect using the probability theory. It represents the probability that an *action* d (i.e., a process deviation) leads to the *outcome* a (i.e., a process activity) while observing the *background* K (i.e., process features). Formally, let $d \in D$ and $a \in A$ be respectively a specific action and outcome, then the causal probability that performing action d results in the outcome a is:

$$P_d(a) = \sum_{k \in K} P(k) \cdot P(a|d\&k). \tag{2}$$

We consider two different models: the first only considers the process instance non-compliance cost $C_{NC}(\mathcal{P}^i)$ while the second also considers additional process contextual attributes (i.e., priority, duration, personnel).

In particular, in our case $P_d(a)$ represents the cost for the deviation $c(a)$ and the different parameters of Eq. (2) are the following:

- $P(k)$ is the probability obtained as the number of traces with feature k over the total number of traces. For example, if k is the feature "SLAs violated", then $P(k) = \frac{\#traces\_violating\_SLAs}{\#traces}$.
- $P(a|d\&k)$ is calculated according to the definition of conditional probability, that is $\frac{P(a \cap d\&k)}{P(d\&k)}$. In particular, $P(a \cap d\&k)$ is the number of traces with deviation d affecting activity a and with features k over the total number of traces, and $P(d\&k)$ is the number of traces with deviation d and features k over the total number of traces.

The assigned costs represent the probabilities of performing certain deviations based on the causality of the features of the incident log, which means the higher the causal probability for a deviation, the more it is a cause of non-compliance.

Linear Regression approach. Statistical regression (Cook and Weisberg, 1982) allows estimating the relationships between a dependent variable Y representing an outcome (i.e., non-compliance cost) and one or more independent variables X representing the observations (i.e., process deviations). This is done by calculating coefficients β_i of dependent variables (i.e., deviations' costs). In particular, multiple linear regression models a linear combination of the independent variables (Bishop and Nasrabadi, 2006). In our case, an observation is a process instance $\mathcal{P}^i$. Y_i is the given non-compliance cost for $\mathcal{P}^i$. X_i is the vector of all the deviations occurring in $\mathcal{P}^i$ such that each element $x_j \in X_i$ is the number of times the deviation x_j appears in the trace $\mathcal{P}^i$. That is, each independent variable (i.e., deviations) is modeled as the number of times it appears in the log. The principle behind this regression design is that redundancy emphasizes the non-compliant behavior of the process (Carmona et al., 2018a). Then, the linear regression model for our problem is:

$$Y_i = \sum_i \beta_i \cdot x_i. \tag{3}$$

The learning task of multiple linear regression is the estimation of the coefficients β_i. Given the proposed linear regression model, such coefficients represent the cost of each deviation x_i (i.e., the weights $c(e_{j,k})$ in Eq. (1)). They are automatically learned given the non-compliance process instances' cost (Y) and the occurrences of each deviation (X). It is applicable when (i) there exists a linear relationship between outcome and independent variables, (ii) the independent variables

are not highly correlated with each other, and (iii) the residuals are normally distributed.

Trees Regression approach. Extra-trees regression (Geurts et al., 2006) is an ensemble technique consisting of the creation of many decision trees, in which each node is a given feature. The extra-trees algorithm randomly samples the features many times in different trees. Then, it chooses the optimal path among all the trees, which minimizes the error expressed in terms of specific metrics as Gini impurity (Breiman et al., 2017) or features entropy (Geurts et al., 2006). The costs of the edges of such a path represent the regression coefficients. For our purposes, each node in the trees represents a deviation, and the number of children depends on the number of times the deviation is repeated. The leaf nodes are the different values of the non-compliance cost. Therefore a path from a root to the leaf node represents a trace and collects the coefficients associated with each deviation to reach the trace cost on the leaf.

The parameters of the extra-trees model are: (i) the number of generated trees, that should be increased until the model performance stabilizes; (ii) the number of features that are randomly sampled for each split point; (iii) the number of samples in a node, with the rationale that a smaller number of samples results in more splits and a more specialized tree. This model is more complex than linear regression in terms of parameters, but more generally applicable due to not being constrained by any assumption.

7. Validation

To validate the proposed model we set up a quantitative benchmark. We need real non-compliance costs associated with IM process instances to populate it. However, by deeply analyzing the literature, we experienced the lack of a standard measure to retrieve such information attached to available activity logs. Therefore, we designed the benchmark using different cost functions taken from the literature, each one covering a different aspect of the IM process considered to compute the non-compliance cost. In the following, we present the design methodology of the benchmark and its implementation and analyses applied to the IM scenario.[3]

Step 1: calculate the cost per trace. The first step of benchmark design is the collection of cost functions representing the non-compliance cost of each trace under different perspectives. While some cost functions apply to any log (being domain agnostic), others may depend on the specific application domain features. We name each cost function as a *ground truth* (GT).

Step 2: estimate non-compliance cost per trace. Although some works estimate such a cost in specific contexts, others do not distinguish between the portion of costs due to non-compliance and the one caused by other factors (e.g., resource expenses). In these cases, we designed a general approach to estimate the non-compliance cost, which considers the cost of non-compliance as part of the trace cost, that is $C_{NC}(\sigma^i) = \alpha \cdot trace\_cost(\sigma^i)$, with $\alpha \in [0, 1]$, sampled with steps of 0.1 to cover each decile of the process execution cost.

Step 3: run experiments. At this stage, we run experiments using all the automatic approaches described in Section 6 and considering each different ground truth. Each approach applied to different ground truths results in different individual costs associated with process deviations.

Step 4: evaluation. We consider the following metrics to validate how much the trace non-compliance costs estimated with the proposed model (Eq. (1)) fit the ones of the ground truths: (i) Mean Absolute Error (MAE) (Bickel and Doksum, 2015), which measures the average absolute difference between estimated values and the actual ones; (ii) Mean Squared Error (MSE) (Bickel and Doksum, 2015), which measures the average squared difference between the estimated values and the actual ones; (iii) Median Absolute Deviation (MAD) (Rousseeuw and Croux, 1993), which measures the variability of a univariate sample of data. More formally, let n be the sample size (i.e., the number of incidents), GT and NC the ground truth and the estimated non-compliance costs respectively. Then, the three error metrics used for the evaluation are calculated as:

$$MAE = \frac{\sum_{i=1}^{n} |GT_i - NC_i|}{n}$$

$$MSE = \frac{\sum_{i=1}^{n} (GT_i - NC_i)^2}{n}$$

$$MAD = |MAD_{GT} - MAD_{NC}|,$$

where $MAD_X = median(|X_i - median(X)|)$. We developed the benchmark by using the reference process model in Fig. 1 and a publicly available dataset containing real data of an IM process from the audit system of the ServiceNow™ (ServiceNow, 2023) platform used by an IT company (Amaral et al., 2019). The log contains 24918 incidents and, for each incident, 33 descriptive features exist related to the IM process (e.g., number of updates during the incidents), incident classification (e.g., categories of the incident and affected services), and incident diagnosis (e.g., causes and impacts).

7.1. Ground truths

The first ground-truth function we consider is based on trace fitness (GT1), as it represents the simplest way to prioritize deviations. Among the approaches defining cost functions for trace costs reviewed in Section 3, we identified four main scientific approaches that consider different aspects of modeling cost functions beyond the fitness-based one. Those functions cover the range of approaches existing in related literature. Two of them evaluate the cost of non-compliance of the traces (GT2, GT3), while the other two calculate the cost of the entire IM process without distinguishing the cost due to non-compliance (GT4, GT5). Since GT4 and GT5 provide the whole cost of process execution, we consider ten variants varying α.

GT1: As the fitness measures the degree of compliance of a trace with a reference process model varying from 0 (non-compliance at all) to 1 (full compliance), we calculate the first ground truth as $1 - fitness$.

GT2 (Kieninger et al., 2013): The cost components are based on losses in revenue, additional expenses, and intangible cost of each incident, and then related to process metrics (e.g., penalty payments). The non-compliance cost is the sum of the *additional* expenses, that is $C_{NC} = t_{add} * p_{add}$, where t_{add} and p_{add} are the working duration and the number of people involved in extra-work for solving the incident.

GT3 (Moura et al., 2006): It estimates the potential loss due to IT Service Level Agreements (SLAs) violations as a function of the duration of each violation. It computes the loss as the rate at which it is instantaneously accumulated at any given time instant, considering the priority of each process. Formally, $Loss = \sum_{\forall i | BP_i \in BP} \sum_j w_i \int_{t \in \tau_j^i} \beta_i(t + \delta_i)dt$, where $Loss$ is the cost due to SLAs violation, BP is the set of the business processes of the organization (e.g., IM process), w is the impact of the process (e.g., incident priority), and β represents the potential revenue rate (e.g., man-hour revenue).

GT4 (Dumas et al., 2013): It describes the cost associated with process resources as the cost of person-hours employed per process instance, thus it is $C(T) = \sum_i t_i * p_i$, where t_i and p_i are respectively the time and the number of people involved in the ith activity of the incident T.

GT5 (Romanosky, 2016): It models a linear regression problem to define the cost in relation to organization revenue (or the number of employees), compromised records, concurrency of incidents, and their impacts. Specifically, the author finds that the relevant factors are the number of employees and the number of compromised records. Specifically, $log(C(T)) = \beta_0 + \beta_1 \cdot log(num\_employees) + \beta_2 \cdot log(records)$, where $\beta_0, \beta_1, \beta_2$ are the regression coefficients and $C(T)$ is the cost of incident T.

[3] https://github.com/Ale96Pa/ComplianceAssessmentSystem/blob/main/test/benchmark/

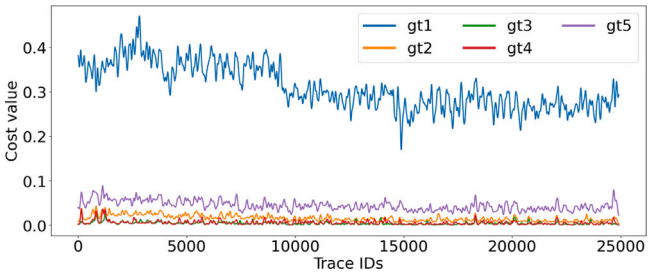


Fig. 4. Comparison between the costs assigned to the incidents according to the different ground truths. The *x*-axis represents the trace IDs and the *y*-axis the trace cost.

7.2. Benchmark analysis

We conducted 92 experiments by running the system with each of the four implemented approaches and the 23 ground truths (GT1, GT2, GT3, ten variants of GT4, and ten variants of GT5), resulting in over 2 million observations. The first research question (RQ1) is: *Is the fitness-based cost (i.e., GT1) a good estimator of the non-compliance cost even considering domain-specific features?* Fig. 4 reports the trace costs calculated with the different ground truths. There is an evident difference between GT1 and all the others. Beyond the cost values, one important aspect is the different trend the fitness has compared with the other ground truths. For example, in the first trace IDs, all the ground truths show an increasing trend, while the fitness has a decreasing one. Similarly, from the trace with ID 5000 to the one with ID 15000, the fitness shows a decreasing trend, while the same is not observed for all the other ground truths that are overall constant. This is a crucial aspect because all the ground truths GT2-GT5 use functions that consider the specific context of the IM process (e.g., incident priority, personnel involved, incident duration). In contrast, the fitness metric only takes into account the control-flow aspect (i.e., differences between log traces and reference model). Its different trend from the other GTs indicates that the fitness does not have the same representativeness of the IM process cost functions. Thus, we can conclude with a negative answer to RQ1. More in detail:

Result 1: *The fitness alone cannot fully capture the context of the process implementation, as its trend differs from all the context-specific ground truths. Contrarily, the ground truths evaluated through domain-specific features have all similar trends, highlighting their common capability to capture the context.*

Because fitness cannot properly express the non-compliance cost, the next research question RQ2 is: *Which automatic approach provides the better estimation for the single ground truths?* We consider the estimated cost trends in Fig. 5 and the error metrics reported in Table 1 (MSE, MAE, and MAD) for each approach and each ground truth. Results highlight the advantages of causal probability and extra-trees models that approximate well the different ground truths. Causal probability, considering the cost as a feature, approximates GT2, GT4, and GT5 better, while extra-trees regression approximates GT3 better. In particular, the causal probability reaches the minimum errors across the benchmark when applied with GT2: MSE is 0.068, MAE 0.170, and MAD 0.078 (bold cell in Table 1). With respect to RQ2, we can conclude:

Result 2: *Extra-trees regression and causal probability considering all the features are comparable to each other: the former is better for GT3 considering all the metrics, and GT4 considering MAD; the latter is better for GT2 and GT5 considering all the metrics, and GT4 considering MSE and MAE. Contrary, linear regression has the highest errors in all the cases considering MSE and MAE.*

Although the previous analysis indicates the error metrics for each individual ground truth, we also examine their distribution among all the experiments to analyze the global best approaches considering all the experiments (Fig. 6(c)). The resulting research question, RQ3, is: *Which automatic approach provides the better estimation considering all the ground truths?* In this analysis, we also consider the fitness-based cost against the other ground truths as it represents the main control-flow metrics used to estimate the degree of non-compliance.

Results are visible in Fig. 6(c). Each box represents the distribution of MSE (Fig. 6(a)), MAE (Fig. 6(b)), and MAD (Fig. 6(c)) of the different approaches. A relevant aspect is that both causal probability and extra-trees regression outperform the prediction of non-compliance cost with respect to the fitness-based approach. Contrary, linear regression performs worst with higher errors. In addition, all the proposed approaches have more compact boxes than fitness-based ones, meaning that the error evaluation is less susceptible to variability. With respect to RQ3 we can conclude:

Result 3: *Causal probability performs better overall considering all the experiments. All the proposed approaches show good performance in terms of non-compliance cost estimation and better capture the non-compliance cost with respect to the only control-flow perspective (i.e., fitness). In addition, our system is more consistent while estimating the non-compliance cost.*

7.3. Accuracy metrics

To investigate the performance of the proposed system further, we provide the evaluation of the accuracy metrics, which are applied to the causal probability approach, considering the cost as a causal feature. To perform this validation, we converted both ground truths and modeled costs into categorical values based on their distribution. For each cost value, the traces costs are labeled as "Low" if they are lower or equal to the median and "High" otherwise. The rationale for choosing the median as the threshold is to avoid bias due to the possible high variability of costs. Then, we compare the labels of ground truth and modeled costs considering the following definitions:

- True Positives (TP) are the observations in which the ground truth is a high cost and our model predicts the same; i.e., the model correctly predicts the critical traces.
- False Positives (FP) are the observations in which the ground truth is a low cost and our model predicts a high cost instead; i.e., the model identifies critical traces that are non-critical instead.
- False Negatives (FN) are the observations in which the ground truth is a high cost and our model predicts a low cost instead; i.e., the model underestimates the criticality of the traces.
- True Negatives (TN) are the observations in which the ground truth is a low cost and our model predicts the same; i.e., the model correctly predicts the non-critical traces.

Fig. 7 reports the confusion matrices of these metrics for each ground truth.

The results in the figure indicate an accuracy of 0.87 for GT2, 0.70 for GT3, 0.67 for GT4, and 0.68 for GT5, resulting in an average accuracy of 0.73. In particular, we are interested in evaluating the trade-off between precision and recall (Bishop and Nasrabadi, 2006), as the former indicates the probability of avoiding false positives and is calculated as $\frac{TP}{TP+FP}$, and the latter indicates the probability of avoiding false negatives and is calculated as $\frac{TP}{TP+FN}$. In our case, the precision values are 0.77, 0.68, 0.64, and 0.62 for GT2, GT3, GT4, and GT5 respectively. Recall values are 0.99, 0.74, 0.73, and 0.75 for GT2, GT3, GT4, and GT5 respectively. This indicates that, on average, the proposed model is able to correctly avoid 67.8% of false positives and 80.3% of false negatives. The precision–recall trade-off indicates that in case of errors the model tends to be conservative, in the sense that

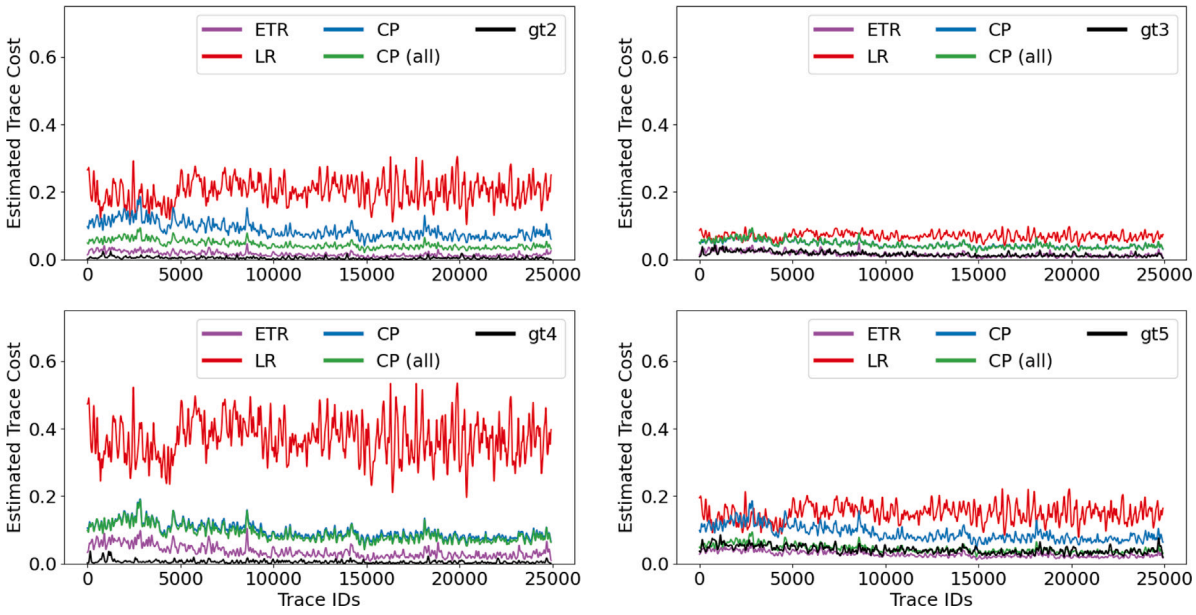
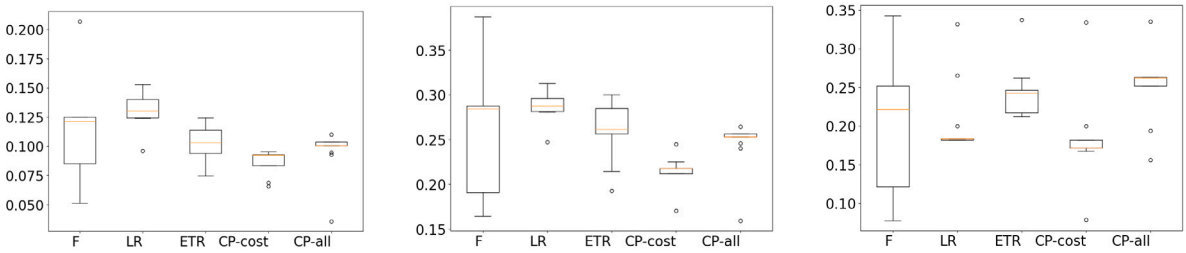


Fig. 5. Comparison between the costs estimated with the different models, i.e., linear regression (LR), extra-tree regression (ETR), causal probability on cost (CP), and causal probability on all attributes (CP-all), for each ground truth (GT2 on top left, GT3 on top right, GT4 on bottom left, GT5 on bottom right).



(a) Mean Squared Error. (b) Mean Absolute Error. (c) Median Absolute Deviation.

Fig. 6. Distribution of the error metrics (a) mean squared error, (b) mean absolute error, and (c) median absolute deviation for each model (F = fitness-based approach, LR = Linear Regression, ETR = Extra-Tree Regression, CP = Causal Probability).

Table 1
Error metrics of modeled approaches compared with each ground truth. The metrics are mse (mean squared error), mae (mean absolute error), and mad (mean absolute deviation).

	GT2			GT3			GT4			GT5		
	MSE	*MAE*	*MAD*	*MSE*	*MAE*	*MAD*	*MSE*	*MAE*	*MAD*	*MSE*	*MAE*	*MAD*
Linear Regression	0.135	0.292	0.332	0.096	0.247	0.265	0.125	0.282	0.183	0.140	0.295	0.181
Extra-Tree Regression	0.114	0.252	0.262	0.074	0.192	0.337	0.094	0.258	0.222	0.112	0.283	0.247
Causal Prob. (cost)	**0.068**	**0.170**	**0.078**	0.094	0.244	0.334	0.092	0.217	0.171	0.083	0.211	0.181
Causal Prob. (all)	0.092	0.240	0.193	0.094	0.245	0.335	0.100	0.253	0.262	0.103	0.256	0.251

it gives more alerts (overestimated costs) rather than underestimating potential dangerous traces.

Result 4: *The probabilistic model has an average accuracy of 0.73. It overestimates 32.2% of the trace costs (false positives) and underestimates 19.7% of them (false negatives). In the presence of errors, the model tends to alert the operator.*

In summary, the quantitative validation results confirm that the proposed model based on causal probability allows for identifying the causes of non-compliance in terms of process deviations and correctly estimating their cost. Focusing on single deviations, the proposed method allows for the characterization of problems from the whole trace granularity to the single deviation occurrences, allowing for a more specific and refined way to prioritize and correct them. Finally, the linear combinations of these deviations, their occurrences, and costs allow for obtaining a fit estimation of the whole trace non-compliance cost.

8. Usage scenario: Incident Management process assessment (ISO 27035)

After discussing the proposed model's quantitative performances, in this section, we illustrate a usage scenario to show the capabilities of the proposed system in a realistic setting. We consider the dataset log (Amaral et al., 2019) and the ISO 27035:2013 as reference process model (ISO/IEC 27035:2013 (E), 2013) (Fig. 1), and the targeted persona is an auditor in charge of assessing the compliance of the IM process through the analysis of its instances in the log.

Once the process auditor loaded the IM log and the reference process model in the system, the first output s/he gets is the analysis of the non-compliance cost (Fig. 8). Fig. 8(a) reports the non-compliance cost evaluated through the causal probability approach (blue) compared with the fitness-based cost (orange) chosen as the simplest way to prioritize deviations and instances. For the sake of presentation and

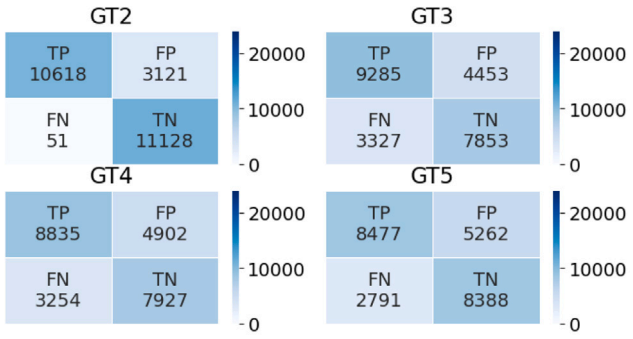


Fig. 7. Confusion matrices of True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN) cost labels.

to allow comparability between fitness and cost model, the traces (x-axis) are sorted by ascending fitness values.[4] In this way, for each trace in the *x*-axis, the two lines show the fitness and non-compliance cost values. This analysis confirms that our cost model is able to differentiate the non-compliance costs based on the type of deviations. This is particularly evident in the incidents from ID 18000 to 20000 (area between red dotted lines in Fig. 8(a)): in that interval, the fitness-based cost is flat, meaning that the traces have all the same non-compliance cost and it would not be possible for the auditor to prioritize some of them. Contrary, our model assigns different costs depending on the different deviations, showing more diverse dynamics of the trend, allowing prioritization of instances to solve.

The auditor then looks at the distribution (box plot) of the non-compliance cost (see Fig. 8(b)). It makes evident that the organization presents good compliance with ISO 27035 (median cost is 0.1 over 1 and lower than 0.2, meaning the overall non-compliance is not particularly critical (Carmona et al., 2018a)). Nonetheless, there are quite a number (4517) of instances covering the remaining non-compliance costs, that should be further investigated as they are the most critical ones. They are highlighted by our proposed cost model but are not visible using standard methods.

The auditor is then interested in exploring the severity distribution for the single instances, and s/he chooses to investigate them further. To this aim, the auditor bins the costs according to *compliance severity* criteria introducing a discrete scale of five values equally distributed in the non-compliance cost range: None, Low, Medium, High, and Critical (see Fig. 9(a)). The bar chart highlights that most of the process instances (21389, corresponding to 85.83%) are affected by "Low" compliance severity, while the "High" and "Critical" traces are respectively 8 and 3 (0.06% of the traces). Thus, 86.15% of the traces present non-compliance issues. However, most of these issues have a minimal impact on the cost of non-compliance. Interestingly, this result shows that focusing on a large number of instances with low non-compliance costs will produce fewer benefits (in terms of non-compliance) than focusing on the medium, high, and critical instances (77 instances versus 21389).

Having identified this interesting subset of instances, the auditor is interested in understanding the causes, so s/he investigates how the non-compliance cost is related to the cost of deviations present in each instance. Fig. 9(b) shows the number of occurrences of each deviation, with their costs estimated through the causal probability approach shown in Table 2.

The deviations are sorted from the most costly to the least costly, which makes it evident to the auditor that *mismatch closure and resolution*, *mismatch resolution*, and *skip detection* are the most critical deviations, with costs 0.97, 0.91, and 0.77 respectively. The most critical deviation is the combination of mismatch closure and mismatch resolution with a cost of 0.97 over 1: it is the case in which, from the incident closure, the process goes back to the "Active" state. It is reasonable that it is the most critical one because it indicates that while the incident response team marked the incident as over, a new threat appeared, forcing them to put the incident state back to "Active". It is interesting to note how the cost associated with the deviations is not directly proportional to their frequency (see Fig. 9(b)). The skipping of detection and activation, and repetition of detection and activation, are the most frequent in the log and they have a cost of 0.77, 0.19, 0.59, and 0.14 respectively.

Table 2
Deviations costs automatically assigned with causal probability approach considering the cost as causal feature.

Deviation	Cost
Mismatch closure and Mismatch resolution	**0.97**
Mismatch resolution	**0.91**
Skip detection	**0.77**
Repeat detection	0.59
Skip resolution	0.53
Skip resolution and Skip detection	0.52
Mismatch closure	0.46
Repeat resolution	0.46

This analysis produces two insights for the auditor: (i) The deviations affecting the detection activity are more severe than the ones affecting the activation. This could be explained by the fact that detection is the very first activity in the reference process model and problems in this phase may result in missing necessary information to process the incidents correctly; (ii) The deviations affecting the activation can be easily repaired as they do not have a critical impact on the IM process. Additionally, this result confirms that the proposed approach assigns costs depending on the process context rather than simply on the frequency of issues, like frequency-based approaches. Thanks to this information, the auditor knows which instances of the IM process are affected by high non-compliance costs (77) and which types of deviations affect them more and must be corrected. This strong reduction allows a more effective intervention and improvement of the IM process. Additionally, by working on them, the auditor will correct even a portion of the non-compliance for less-costly instances, getting a second-level effect on the overall improvement of the IM process compliance.

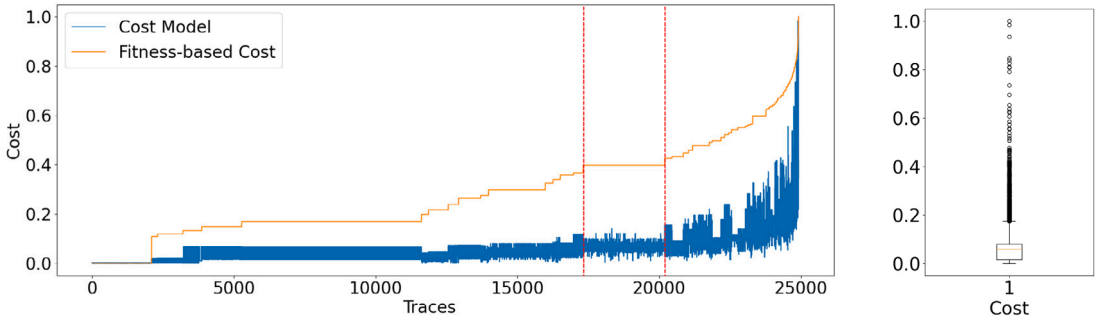
Notice that gaining the same information through a manually performed assessment would have been significantly more time and resource-consuming (i.e., interviews, questionnaires). For example, considering only the trace costs without the proposed model, an auditor would have analyzed the most critical traces (e.g., the worst 10). It corresponds to the analysis of 310 deviations with an average of 31 deviations per trace. This is because the analysis performed by trace does not consider common deviations among the traces. Contrarily, the proposed system prioritizes the deviations, highlighting, in this case, that fixing the three most critical deviations would reduce 53.5% of the non-compliance cost.

In conclusion, the system helps the auditor prioritize mitigation actions. S/he identifies, based on these results, that the most urgent investment is in the response team training because the high cost of mismatch closure and resolution highlights the need to better recognize the closure conditions of an incident. In addition, the high cost of skip detection deviation highlights the need to train users better to detect incidents promptly.

9. Discussion and limitations

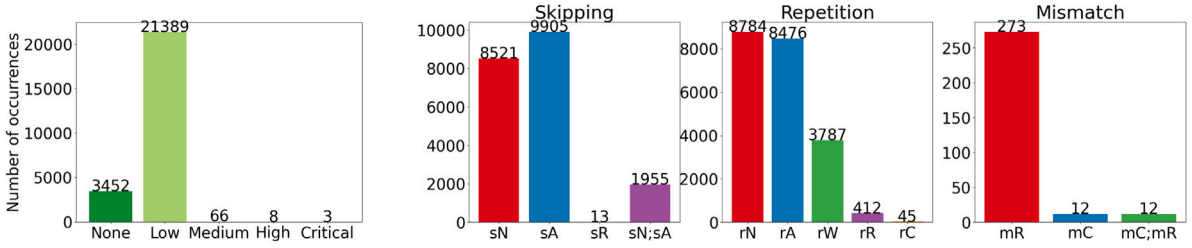
This paper proposed a novel approach to support the auditor during the IM process compliance assessment. We addressed three main

[4] Without loss of generality, it is always possible to retrieve the original trace IDs from the ones sorted by fitness.



(a) (b)

Fig. 8. (a) Non-compliance cost trend and (b) non-compliance cost distribution. The area between the two dotted lines highlights the flatness of fitness-based cost with respect to the proposed cost model.



(a) Compliance severity distribution. (b) Occurrences per deviation type.

Fig. 9. (a) Deviations analysis per compliance severity and (b) deviations analysis per deviation type (s: skip, r: repetition, m: mismatch). The following notation is used for the different Incident Management activities (N: detection, A: activation, W: awaiting for third-party information, R: resolution, C: closure). The *y*-axis represents the number of occurrences of the different deviations, also reported at the top of each bar.

Table 3
Comparison of compliance assessment methodologies. The columns denote the presence of quantitative evaluation, automation, and if it is agnostic to security standards. We further specify the associated compliance metric and the methodology used (ND = Not Defined).

	Quantitative assessment	Automated approach	Standard agnostic	Compliance metric	Methodology
Arsac et al. (2011)	×	✓	✓	ND	model checking
Ly et al. (2012)	×	✓	✓	ND	rule-based
Sarkar and Saren (2016)	✓	×	✓	manufacturing cost	mathematical model
Kabaale et al. (2018)	×	×	×	ND	ontology
Glogovac et al. (2019)	✓	×	✓	quality cost	mathematical model
Angelini et al. (2020)	×	×	✓	coverage validation	attack graph
Shinde and Kulkarni (2021)	×	×	×	ND	user study
van der Kleij et al. (2022)	×	×	×	ND	user study
He et al. (2022)	×	×	×	ND	framework
Alfaadhel et al. (2023)	✓	×	×	compliance score	questionnaires
Mouratidis et al. (2023)	×	✓	✓	ND	framework
Ghanem et al. (2023)	×	×	✓	ND	rule-based
Proposed approach	✓	✓	✓	**non-compliance cost**	**trace alignment**

research questions. The goal of the first one is to understand whether the fitness metric, which represents the current state-of-the-art metric for process compliance, is adequate for estimating the non-compliance cost of the IM process. The performed validation showed that the fitness metric cannot fully capture the context of the IM process implementation because it does not consider specific incident features (e.g., incident priority). This implies that more advanced methodologies are necessary to estimate the non-compliance cost. To this aim, the second research question analyzes which is the best automatic approach, among the ones proposed, to estimate non-compliance costs. The results showed that there is no best approach, but Extra-Tree regression and Causal Probability have comparable performance. Consequently, the third research question investigates the overall performance of the automatic approaches in a benchmark validation, which showed that the Causal Probability approach outperforms the others when considering the entire set of experiments, with an average accuracy of 0.7.

We report in Table 3 the comparison of the proposed work with the current state-of-the-art of IM compliance process assessment. It shows that most of the literature focuses on qualitative assessments by contributing frameworks (Mouratidis et al., 2023; He et al., 2022), rule-based approaches (Ly et al., 2012; Ghanem et al., 2023), and user studies (Shinde and Kulkarni, 2021; van der Kleij et al., 2022). This hinders the possibility of measuring compliance through suitable quantitative metrics (ND labels in Table 3 stand for Not Defined). In contrast, the few works proposing quantitative assessment either provide a metric not measuring the IM process compliance but rather focusing on manufacturing costs (Glogovac et al., 2019; Sarkar and

Saren, 2016) or require manual effort to analyze checklist-based security standards (Angelini et al., 2020; Alfaadhel et al., 2023). Finally, five works are not standard agnostic and therefore can only be applied for compliance with a specific security standard (e.g., Alfaadhel et al., 2023 support only the ECC standard). To the best of the authors' knowledge, the proposed approach represents the first contribution that allows for quantitatively assessing, supported by automatic computation, the compliance of the IM process with a reference process model, identifying the non-compliance causes in terms of process deviations, and prioritizing them by assigning costs. Indeed, it is the only one leveraging the methodology of trace alignment, a process mining technique, and extending it through a cost model for better accuracy. Although we modeled the approach, its quantitative validation, and usage scenario for the IM process, we believe it can be easily generalized to any other cybersecurity governance process by suitably formatting a reference process model and the trace log containing its execution.

Further considerations arise from this work in the form of limitations, reported in the following:

Modeling reference processes. Modeling a process model from the security standards according to business process modeling notations (e.g., Petri Net (Petri, 1966), BPMN (White, 2004), WorkFlow Net (Salimifard and Wright, 2001)) may still require some effort when not supported by existing standards. This is because the designer should have technical, business, knowledge management, and social competencies (Sonteya and Seymour, 2012).

We consider the reference process model as an input of the system. Therefore the difficulty of its modeling is out of the scope of this paper. On the other hand, we observe that the processes described in the security standards are typically general and therefore consist of a simple sequence of activities (e.g., planning, detection, assessment, response, and closure for IM). In such cases, the process modeling is easier and can be achieved with intuitive tools such as WoPeD[5] without requiring high expertise.

Log quality and requirements. The second element that we consider as input of the system is the IM process log. We consider the log quality only from the structural perspective, meaning that any valid log must contain the trace IDs, the phases in each trace, and the timestamp of each phase (Van Der Aalst, 2013). In addition, in the proposed system the log must also contain the non-compliance cost of each trace. As it is typically an attribute difficult to retrieve in the logs, we proposed five existing cost functions to estimate them. We do not consider the quality of the logged data instead (e.g., the presence of inconsistent values) and how it impacts the assessment. According to the literature on log data quality (Kherbouche et al., 2016; Bose et al., 2013), three possible scenarios are possible during the IM process: (i) missing data (i.e., the operators or the automatic data collection system do not fill the log entries), (ii) wrong data (i.e., the operators or the automatic data collection system insert a wrong value for a feature of the log), and (iii) incomplete data (i.e., the operators or the automatic data collection system do not fill all the entries/features of the same incident) (Palma et al., 2024). In the proposed system, missing and incomplete data results in a smaller training set for the automated cost assignment approaches. A typical solution for this problem is adding the constraint of a sufficient number of entries necessary to train the approaches (e.g., ten). In contrast, wrong data may affect the accuracy of the automated cost assignment approaches. To address this problem, a possible solution is to introduce noise in the benchmark to evaluate how much the approaches are robust to the introduced noise. Then, the auditor is more informed about the robustness of the system and can weigh her decision accordingly (e.g., if s/he knows that the log may contain many wrong values, s/he can discard the result in case of low robustness). Applying those mitigations can limit the potential effects of log quality on the analysis results. We leave a deeper investigation of this analysis for future work.

Applicability to different IM logs. The proposed compliance assessment system currently handles only a single IM log for experimental validation. This is because incident data are typically not shared publicly, making them difficult to obtain. To the best of the authors' knowledge, this is the only dataset for IM that includes both process and incident features. However, since this work is public, readers can download the code and run the system with additional (private) logs. It is important to note that different datasets only impact the performance of the automated cost assignment techniques (i.e., regression and probabilistic). The selected technique may vary across datasets, and evaluating the most performant approach for all the datasets is out of the scope of this paper. A possible solution to address this problem is the generation of synthetic logs inferred from the real one to study how the different approaches vary across controlled variations of the dataset (Palma et al., 2024).

10. Conclusions

This paper addressed the problem of compliance assessment for the Incident Management process. In doing so, it contributes a novel deviation taxonomy, its related cost model, and a compliance assessment system to support an auditor in assessing (i) the IM process compliance with a reference process model and (ii) the prioritization of instances to mitigate based on the deviation costs and presence. This approach tested quantitatively and qualitatively, supports the auditor in making faster and more effective the identification of more costly errors during the IM process, correcting them, and estimating the corrections' impact on the rest of the logged instances.

The experimental evaluation showed four important results from this study. The first outcome indicates that the fitness metric does not account for specific incident characteristics (e.g., impact, personnel), necessitating more advanced methodologies for accurate non-compliance cost estimation. In contrast, the proposed cost model addresses this challenge. The second result reveals no single best method for estimating the ground truths, but Extra-Tree regression and Causal Probability approaches show comparable and good performance when considering each single ground truth. The third and fourth findings assess the performance of the automatic approaches over all the ground truths. The former demonstrates that the Causal Probability outperforms others in terms of aggregated error metrics, while the latter shows its better performance in terms of achieved accuracy. The usage scenario demonstrates the benefits of using the proposed system to assess compliance with ISO 27035 and prioritize mitigation strategies accordingly. The acquired findings showed the advancement of the proposed system in the current literature as it quantitatively assesses IM process compliance with a reference model through a novel cost model based on trace alignment. It uses automated approaches for cost computation that show good performance for estimating the ground truths and identifying and prioritizing non-compliance causes through process deviations.

In future works, we plan to work on the following three research directions. The first one concerns quantitative validation. Although correctly covering the main literature approaches, it uses only one dataset since real IM process logs are difficult to retrieve. For this purpose, we plan to involve companies to expand the validation and propose a complete benchmark useful for any future validation activities (Palma et al., 2024). The second research direction is focused on the interpretation of the assessment results. We plan to provide more control to the auditor during the process assessment by leveraging Visual Analytics techniques (Keim et al., 2008; Palma and Angelini, 2024). The last research direction deals with the dynamic estimation of non-compliance costs by leveraging the context-aware trace alignment (Acitelli et al., 2022).

[5] https://woped.dhbw-karlsruhe.de/

CRediT authorship contribution statement

Alessandro Palma: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Giacomo Acitelli:** Writing – review & editing, Writing – original draft, Investigation, Data curation, Conceptualization. **Andrea Marrella:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Formal analysis, Data curation, Conceptualization. **Silvia Bonomi:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marco Angelini:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data and code are openly available and linked in the paper.

Acknowledgments

This work has been partially supported by project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU. The work of G. Acitelli and A. Marrella has been supported by the Sapienza Project FOND-AIBPM, the PRIN 2022 project MOTOWN and the PNRR MUR project PE0000013-FAIR.

References

Accorsi, R., 2009. Safe-keeping digital evidence with secure logging protocols: State of the art and challenges. In: 2009 Fifth International Conference on IT Security Incident Management and IT Forensics. IEEE, pp. 94–110. http://dx.doi.org/10.1109/IMF.2009.18.

Accorsi, R., Stocker, T., 2012. On the exploitation of process mining for security audits: the conformance checking case. In: Proceedings of the 27th Annual ACM Symposium on Applied Computing. pp. 1709–1716. http://dx.doi.org/10.1145/2245276.2232051.

Acitelli, G., Angelini, M., Bonomi, S., Maggi, F.M., Marrella, A., Palma, A., 2022. Context-aware trace alignment with automated planning. In: 2022 4th International Conference on Process Mining. ICPM, IEEE, pp. 104–111. http://dx.doi.org/10.1109/ICPM57379.2022.9980649.

Adriansyah, A., Sidorova, N., van Dongen, B.F., 2011. Cost-based fitness in conformance checking. In: Proc. of Int. Conf. on Application of Concurrency To System Design. IEEE, pp. 57–66. http://dx.doi.org/10.1109/ACSD.2011.19.

Aldasoro, I., Gambacorta, L., Giudici, P., Leach, T., 2022. The drivers of cyber risk. J. Financial Stab. 60, 100989. http://dx.doi.org/10.1016/j.jfs.2022.100989.

Alfaadhel, A., Almomani, I., Ahmed, M., 2023. Risk-based cybersecurity compliance assessment system (RC2AS). Appl. Sci. 13 (10), http://dx.doi.org/10.3390/app13106145.

Ali, R.F., Dominic, P.D.D., Ali, S.E.A., Rehman, M., Sohail, A., 2021. Information security behavior and information security policy compliance: A systematic literature review for identifying the transformation process from noncompliance to compliance. Appl. Sci. 11 (8), http://dx.doi.org/10.3390/app11083383.

Amaral, C.A.L., Fantinato, M., Reijers, H.A., Peres, S.M., 2019. Enhancing completion time prediction through attribute selection. In: Ziemba, E. (Ed.), Information Technology for Management: Emerging Research and Applications. Springer International Publishing, Cham, pp. 3–23. http://dx.doi.org/10.1007/978-3-030-15154-6_1.

Angelini, M., Bonomi, S., Ciccotelli, C., Palma, A., 2020. Toward a context-aware methodology for information security governance assessment validation. In: International Workshop on Cyber-Physical Security for Critical Infrastructures Protection. Springer, pp. 171–187. http://dx.doi.org/10.1007/978-3-030-69781-5_12.

Angelini, M., Bonomi, S., Palma, A., 2022. A methodology to support automatic cyber risk assessment review. arXiv preprint arXiv:2207.03269.

Arsac, W., Compagna, L., Pellegrino, G., Ponta, S.E., 2011. Security validation of business processes via model-checking. In: International Symposium on Engineering Secure Software and Systems. Springer, pp. 29–42. http://dx.doi.org/10.1007/978-3-642-19125-1_3.

Augusto, A., Conforti, R., Dumas, M., La Rosa, M., Maggi, F.M., Marrella, A., Mecella, M., Soo, A., 2018. Automated discovery of process models from event logs: review and benchmark. IEEE TKDE 31 (4), 686–705. http://dx.doi.org/10.1109/TKDE.2018.2841877.

Bernardi, S., Domínguez, J.L., Gómez, A., Joubert, C., Merseguer, J., Perez-Palacin, D., Requeno, J.I., Romeu, A., 2018. A systematic approach for performance assessment using process mining: An industrial experience report. Empir. Softw. Eng. 23 (6), 3394–3441. http://dx.doi.org/10.1007/s10664-018-9606-9.

Berti, A., van Zelst, S.J., van der Aalst, W.M.P., 2019. Process mining for python (PM4Py): Bridging the gap between process- and data science. In: ICPMD 2019, ICPM Demo Track 2019. In: CEUR workshop proceedings, vol. 2374, RWTH Aachen, Aachen, Germany, pp. 13–16, 1st International Conference on Process Mining, Aachen (Germany), 24 Jun 2019 - 24 Jun 2019.

Bertrand, Y., De Weerdt, J., Serral, E., 2023. A novel multi-perspective trace clustering technique for IoT-enhanced processes: A case study in smart manufacturing. In: Di Francescomarino, C., Burattin, A., Janiesch, C., Sadiq, S. (Eds.), Business Process Management. Springer Nature, Switzerland, Cham, pp. 395–412. http://dx.doi.org/10.1007/978-3-031-41620-0_23.

Bickel, P.J., Doksum, K.A., 2015. Mathematical Statistics: Basic Ideas and Selected Topics, Volumes I-II Package. Chapman and Hall/CRC.

Bishop, C.M., Nasrabadi, N.M., 2006. Pattern Recognition and Machine Learning, vol. 4, Springer, URL http://research.microsoft.com/cmbishop/PRML.

Bose, R.J.C., Mans, R.S., Van Der Aalst, W.M., 2013. Wanna improve process mining results? In: 2013 IEEE Symposium on Computational Intelligence and Data Mining. CIDM, pp. 127–134. http://dx.doi.org/10.1109/CIDM.2013.6597227.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 2017. Classification and Regression Trees. Routledge, http://dx.doi.org/10.1002/widm.8.

Carmona, J., van Dongen, B.F., Solti, A., Weidlich, M., 2018a. Conformance Checking - Relating Processes and Models. Springer, http://dx.doi.org/10.1007/978-3-319-99414-7.

Caron, F., Vanthienen, J., Baesens, B., 2012. Comprehensive rule-based compliance checking and risk management with process mining. http://dx.doi.org/10.1016/j.dss.2012.12.012.

Cook, R.D., Weisberg, S., 1982. Criticism and influence analysis in regression. Sociol. Methodol. 13, 313–361. http://dx.doi.org/10.2307/270724.

De Leoni, M., Marrella, A., 2017. Aligning real process executions and prescriptive process models through automated planning. Expert Syst. Appl. 82, 162–183. http://dx.doi.org/10.1016/j.eswa.2017.03.047.

De Leoni, M., Munoz-Gama, J., Carmona, J., Van Der Aalst, W.M.P., 2014. Decomposing alignment-based conformance checking of data-aware process models. In: Meersman, R., Panetto, H., Dillon, T., Missikoff, M., Liu, L., Pastor, O., Cuzzocrea, A., Sellis, T. (Eds.), In: On the Move To Meaningful Internet Systems: OTM 2014 Conferences, vol. 8841, Springer, Berlin, Heidelberg, pp. 3–20. http://dx.doi.org/10.1007/978-3-662-45563-0_1, Series Title: Lecture Notes in Computer Science.

De Weerdt, J., Vanden Broucke, S.K., Vanthienen, J., Baesens, B., 2012. Leveraging process discovery with trace clustering and text mining for intelligent analysis of incident management processes. In: 2012 IEEE Congress on Evolutionary Computation. IEEE, pp. 1–8. http://dx.doi.org/10.2139/ssrn.2165170.

Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A., et al., 2013. Fundamentals of Business Process Management, vol. 1, Springer, http://dx.doi.org/10.1007/978-3-662-56509-4.

El Kharbili, M., 2012. Business process regulatory compliance management solution frameworks: A comparative evaluation. In: Proceedings of the Eighth Asia-Pacific Conference on Conceptual Modelling-Volume 130. pp. 23–32. http://dx.doi.org/10.5555/2523782.2523786.

ENISA, 2010. Good practice guide for incident management. URL https://www.enisa.europa.eu/publications/good-practice-guide-for-incident-management,

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach. Learn. 63 (1), 3–42. http://dx.doi.org/10.1007/s10994-006-6226-1.

Ghanem, M.C., Chen, T.M., Ferrag, M.A., Kettouche, M.E., 2023. ESASCF: Expertise extraction, generalization and reply framework for optimized automation of network security compliance. IEEE Access 11, 129840–129853. http://dx.doi.org/10.1109/ACCESS.2023.3332834.

Glogovac, M., Filipovic, J., Zivkovic, N., Jeremic, V., 2019. A Model for Prioritization of Improvement Opportunities Based on Quality Costs in the Process Interdependency Context. Eng. Econ. 30 (3), 278–293. http://dx.doi.org/10.5755/j01.ee.30.3.14657, Number: 3.

Gohil, F., Kumar, M.V., 2019. Ticketing system. Int. J. Trend Sci. Res. Dev. 3 (4), 155–156. http://dx.doi.org/10.31142/ijtsrd23603.

González-Granadillo, G., González-Zarzosa, S., Diaz, R., 2021. Security information and event management (SIEM): Analysis, trends, and usage in critical infrastructures. Sensors 21 (14), http://dx.doi.org/10.3390/s21144759.

He, Y., Zamani, E.D., Lloyd, S., Luo, C., 2022. Agile incident response (AIR): Improving the incident response process in healthcare. Int. J. Inf. Manage. 62, 102435. http://dx.doi.org/10.1016/j.ijinfomgt.2021.102435.

ISO 19600:2014 (E), 2014. Compliance management systems — Guidelines, vol. 2014, International Organization for Standardization, Geneva, CH.

ISO 37301:2021 (E), 2021. Compliance Management Systems — Requirements with Guidance for Use. Standard, vol. 2021, International Organization for Standardization, Geneva, CH.

ISO/IEC 27035:2013 (E), 2013. Part 1: Principles of Incident Management; Part 2: Guidelines to Plan and Prepare for Incident Response; Part 3: Guidelines for ICT Incident Response Operations. International Organization for Standardization.

ISO/TC 9001, 2014. Quality Management Systems. Standard, vol. 2014, International Organization for Standardization, Geneva, CH.

ITILv4, 2019. Information Technology Infrastructure Library. Standard, vol. 2019, Axelos, UK.

Jadhav, A., Kaur, M., Akter, F., 2022. Evolution of software development effort and cost estimation techniques: five decades study using automated text mining approach. Math. Probl. Eng. 2022, 1–17. http://dx.doi.org/10.1155/2022/5782587.

Kabaale, E., Wen, L., Wang, Z., Rout, T., 2018. Ensuring conformance to process standards through formal verification. In: Stamelos, I., O'Connor, R.V., Rout, T., Dorling, A. (Eds.), In: Software Process Improvement and Capability Determination, vol. 918, Springer International Publishing, Cham, pp. 248–262. http://dx.doi.org/10.1007/978-3-030-00623-5_17, Series Title: Communications in Computer and Information Science.

Kazmer, D., Peterson, A.M., Masato, D., Colon, A.R., Krantz, J., 2023. Strategic cost and sustainability analyses of injection molding and material extrusion additive manufacturing. Polym. Eng. Sci. 63 (3), 943–958. http://dx.doi.org/10.1002/pen.26256.

Keim, D.A., Mansmann, F., Schneidewind, J., Thomas, J., Ziegler, H., 2008. Visual analytics: Scope and challenges. Visual Data Mining: Theory, Techniques and Tools for Visual Analytics. Springer, Berlin, Heidelberg, pp. 76–90. http://dx.doi.org/10.1007/978-3-540-71080-6_6.

Kherbouche, M.O., Laga, N., Masse, P.-A., 2016. Towards a better assessment of event logs quality. In: 2016 IEEE Symposium Series on Computational Intelligence. SSCI, pp. 1–8. http://dx.doi.org/10.1109/SSCI.2016.7849946.

Kieninger, A., Berghoff, F., Fromm, H., Satzger, G., 2013. Simulation-Based Quantification of Business Impacts Caused by Service Incidents. In: Van Der Aalst, W., Mylopoulos, J., Rosemann, M., Shaw, M.J., Szyperski, C., Falcão e Cunha, J.a., Snene, M., Nóvoa, H. (Eds.), In: Exploring Services Science, vol. 143, Springer, Berlin, Heidelberg, pp. 170–185. http://dx.doi.org/10.1007/978-3-642-36356-6_13, Series Title: Lecture Notes in Business Information Processing.

Kothandapani, H.P., 2023. Applications of robotic process automation in quantitative risk assessment in financial institutions. Int. J. Bus. Intell. Big Data Anal. 6 (1), 40–52, URL https://research.tensorgate.org/index.php/IJBIBDA/article/view/80.

Liu, Y., Muller, S., Xu, K., 2007. A static compliance-checking framework for business process models. IBM Syst. J. 46 (2), 335–361. http://dx.doi.org/10.1147/sj.462.0335.

Ly, L.T., Rinderle-Ma, S., Göser, K., Dadam, P., 2012. On enabling integrated process compliance with semantic constraints in process management systems. Inf. Syst. Front. 14 (2), 195–219. http://dx.doi.org/10.1007/s10796-009-9185-9.

Madigan, E.M., Petrulich, C., Motuk, K., 2004. The cost of non-compliance: When policies fail. In: Proceedings of the 32nd Annual ACM SIGUCCS Conference on User Services. SIGUCCS '04, Association for Computing Machinery, New York, NY, USA, pp. 47–51. http://dx.doi.org/10.1145/1027802.1027815.

Márquez-Chamorro, A.E., Resinas, M., Ruiz-Cortés, A., 2017. Predictive monitoring of business processes: A survey. IEEE Trans. Serv. Comput. 11 (6), 962–977. http://dx.doi.org/10.1109/TSC.2017.2772256.

Moura, A., Sauve, J., Jornada, J., Radziuk, E., 2006. A Quantitative Approach to IT Investment Allocation to Improve Business Results. In: Seventh IEEE International Workshop on Policies for Distributed Systems and Networks. POLICY'06, IEEE, London, ON, Canada, pp. 87–95. http://dx.doi.org/10.3905/jpm.2018.44.2.156.

Mouratidis, H., Islam, S., Santos-Olmo, A., Sanchez, L.E., Ismail, U.M., 2023. Modelling language for cyber security incident handling for critical infrastructures. Comput. Secur. 128, 103139. http://dx.doi.org/10.1016/j.cose.2023.103139.

Naseer, A., Naseer, H., Ahmad, A., Maynard, S.B., Siddiqui, A.M., 2023. Moving towards agile cybersecurity incident response: A case study exploring the enabling role of big data analytics-embedded dynamic capabilities. Comput. Secur. 135, 103525. http://dx.doi.org/10.1016/j.cose.2023.103525.

National Institute of Standards and Technology, 2021. NIST special publication 800-61, revision 2, computer security incident handling guide. URL https://www.nist.gov/privacy-framework/nist-sp-800-61.

Nguyen, H., Dumas, M., Rosa, M.L., Maggi, F.M., Suriadi, S., 2014. Mining business process deviance: A quest for accuracy. In: On the Move To Meaningful Internet Systems: OTM 2014 Conferences - Confederated International Conferences: CoopIS, and ODBASE 2014, Amantea, Italy, October 27-31, 2014, Proceedings. Springer, pp. 436–445. http://dx.doi.org/10.1007/978-3-662-45563-0_25.

Palma, A., Angelini, M., 2024. Visually Supporting the Assessment of the Incident Management Process. In: El-Assady, M., Schulz, H.-J. (Eds.), EuroVis Workshop on Visual Analytics (EuroVA). The Eurographics Association, http://dx.doi.org/10.2312/eurova.20241116.

Palma, A., Bartoloni, N., Angelini, M., 2024. BenchIMP: A benchmark for quantitative evaluation of the incident management process assessment. In: Proceedings of the 19th International Conference on Availability, Reliability and Security. ARES '24, Association for Computing Machinery, New York, NY, USA, http://dx.doi.org/10.1145/3664476.3664504.

Pascual, R., Knights, P., Louit, D., Castillo, G.D., 2009. Business-oriented prioritization: A novel graphical technique. http://dx.doi.org/10.1016/j.ress.2009.01.013.

Petri, C.A., 1966. Communication with Automata. Tech. rep., Hamburg University, URL https://api.semanticscholar.org/CorpusID:8420535.

Pramanik, N., Roy, U., Sudarsan, R., Sriram, R., Lyons, K., 2005. A generic deviation-based approach for synthesis of tolerances. IEEE Trans. Autom. Sci. Eng. 2 (4), 358–368. http://dx.doi.org/10.1109/TASE.2005.853584, Conference Name: IEEE Transactions on Automation Science and Engineering.

Romanosky, S., 2016. Examining the costs and causes of cyber incidents. J. Cybersecur. tyw001. http://dx.doi.org/10.1093/cybsec/tyw001.

Rousseeuw, P.J., Croux, C., 1993. Alternatives to the median absolute deviation. J. Amer. Statist. Assoc. 88 (424), 1273–1283. http://dx.doi.org/10.2307/2291267.

Salimifard, K., Wright, M., 2001. Petri net-based modelling of workflow systems: An overview. European J. Oper. Res. 134 (3), 664–676. http://dx.doi.org/10.1016/S0377-2217(00)00292-7.

Santos, A.C., Willumsen, J., Meheus, F., Ilbawi, A., Bull, F.C., 2023. The cost of inaction on physical inactivity to public health-care systems: A population-attributable fraction analysis. Lancet Global Health 11 (1), e32–e39. http://dx.doi.org/10.1016/S2214-109X(22)00464-8.

Sarkar, B., Saren, S., 2016. Product inspection policy for an imperfect production system with inspection errors and warranty cost | Elsevier Enhanced Reader. http://dx.doi.org/10.1016/j.ejor.2015.06.021.

ServiceNow, 2023. ServiceNow-TM. Tech. rep., Gildesoft, USA, URL https://docs.servicenow.com/.

Shinde, N., Kulkarni, P., 2021. Cyber incident response and planning: A flexible approach. Comput. Fraud Secur. 2021 (1), 14–19. http://dx.doi.org/10.1016/S1361-3723(21)00009-9.

Silalahi, S., Yuhana, U.L., Ahmad, T., Studiawan, H., 2022. A survey on process mining for security. In: 2022 International Seminar on Application for Technology of Information and Communication (ISemantic). pp. 1–6. http://dx.doi.org/10.1109/iSemantic55962.2022.9920473.

Siponen, M., Willison, R., 2009. Information security management standards: Problems and solutions. Inf. Manag. 46 (5), 267–270. http://dx.doi.org/10.1016/j.im.2008.12.007.

Skyrms, B., 1982. Causal decision theory. J. Phil. 79 (11), 695–711. http://dx.doi.org/10.2307/2026547.

Solarwind, 2021. Solarwind. Standard, vol. 2021, SolarWinds Corporation, USA, URL https://www.solarwinds.com/.

Sonteya, T., Seymour, L.F., 2012. Towards an understanding of the business process analyst: An analysis of competencies. J. Inf. Technol. Educ.: Res. 11 (1), 43–63. http://dx.doi.org/10.28945/1568.

Vaarandi, R., 2005. Tools and techniques for event log analysis. URL https://api.semanticscholar.org/CorpusID:13360648.

Van Der Aalst, W., 2012. Process mining. Commun. ACM 55 (8), 76–83. http://dx.doi.org/10.1145/2240236.2240257.

Van Der Aalst, W.M., 2013. Business process management: A comprehensive survey. Int. Sch. Res. Notices 2013, http://dx.doi.org/10.1155/2013/507984.

Van Der Aalst, W.M.P., 2016. Process Mining - Data Science in Action, Second Edition Springer, http://dx.doi.org/10.1007/978-3-662-49851-4.

van der Kleij, R., Schraagen, J.M., Cadet, B., Young, H., 2022. Developing decision support for cybersecurity threat and incident managers. Comput. Secur. 113, 102535. http://dx.doi.org/10.1016/j.cose.2021.102535.

Vanden Broucke, S.K., De Weerdt, J., Vanthienen, J., Baesens, B., 2013. A comprehensive benchmarking framework (CoBeFra) for conformance analysis between procedural process models and event logs in ProM. In: 2013 IEEE Symposium on Computational Intelligence and Data Mining. CIDM, pp. 254–261. http://dx.doi.org/10.1109/CIDM.2013.6597244.

Vanounou, T., Pratt, W., Fischer, J.E., Jr, C.M.V., Callery, M.P., 2007. Deviation-Based Cost Modeling: A Novel Model to Evaluate the Clinical and Economic Impact of Clinical Pathways. http://dx.doi.org/10.1016/j.jamcollsurg.2007.01.025.

Varela-Vaca, A., Gasca, R.M., Jimenez-Ramirez, A., 2011. A model-driven engineering approach with diagnosis of non-conformance of security objectives in business process models. In: 2011 Fifth International Conference on Research Challenges in Information Science. pp. 1–6. http://dx.doi.org/10.1109/RCIS.2011.6006844, ISSN: 2151-1357.

Waspada, I., Sarno, R., Astuti, E.S., Prasetyo, H.N., Budiraharjo, R., 2022. Graph-based token replay for online conformance checking. IEEE Access 10, 102737–102752. http://dx.doi.org/10.1109/ACCESS.2022.3208098.

White, S.A., 2004. Introduction to BPMN. Ibm Coop. 2, URL http://yoann.nogues.free.fr/IMG/pdf/07-04_WP_Intro_to_BPMN_-_White-2.pdf.

Wibawa, N.M.S., Ramantoko, G., 2022. Business process analysis of cloud incident management service with activity assignment: A case of PT. XYZ. J. Bus. Manag. Account. 12 (1), 51–80. http://dx.doi.org/10.32890/jbma2022.12.1.3.

Alessandro Palma is a Ph.D. student in Engineering in Computer Science in the Department of Computer, Control, and Management Engineering "Antonio Ruberti" at Sapienza University of Rome, Italy. He received the M.Sc. in Engineering in Computer Science from Sapienza University of Rome. He worked as a research fellow with CINI Cybersecurity National Laboratory on OSINT topics. His main research interests focus on support for correct and quantifiable cyber risk assessment, researching attack modeling and attack graphs, and providing automatic support for security governance processes analysis. More information available at https://sites.google.com/diag.uniroma1.it/palma

Giacomo Acitelli is a Ph.D. student in Engineering in Computer Science at Sapienza University of Rome, Italy. His research interests span from Business Process Management and Process Mining to Reasoning About Actions in Artificial Intelligence. He is currently involved in Italian and European projects on developing advanced AI-based solutions for Process Mining in many real-world domains, including Cybersecurity.

Andrea Marrella is an Associate Professor of Engineering in Computer Science at DIAG. His research work concerns the application of techniques for Reasoning About Actions, including Automated Planning, to problems arising in Business Process Management, Robotic Process Automation, and Process Mining. Andrea has co-authored over 100 scientific publications in major outlets in the AI and information systems areas, winning a Best Paper Award at CAiSE 2017. Andrea is an Editorial Board Member of ACM Computing Surveys and ACM Journal of Data and Information Quality. Among his recent scientific appointments, he has been the PC Chair of the RPA track of BPM 2022 and will be the General Chair of ICPM 2023 and PC Co-Chair of BPM 2024. In 2022, he co-authored the Manifesto on AI-Augmented BPM, presenting the vision to make processes more adaptable, proactive and explainable through AI. In 2023, he was appointed to organize and chair the AAAI 2023 Bridge Program on AI and BPM. From 2022, he coordinates the working group on AI and BPM of the AIxIA community. Since 2021, he is the Local PI of the H2020 project DataCloud, focusing on empowering process mining with AI to develop a new breed of Big Data pipeline discovery solutions.

Silvia Bonomi is Associate Professor at Sapienza University of Rome. She got a joint Ph.D. in Computer Science and Computer Engineering. She is member of Cyber Intelligence and Information Security research group (CIS-Sapienza) of the Department of Computer, Control, and Management Engineering "Antonio Ruberti" since 2012 and of the Midlab research group since 2006. She is also member of the CINI Cyber Security National Laboratory. She currently leads the Distributed Systems research group at DIAG Sapienza. She has been and is currently involved in several National (PNRR PE7, PNRR CN2, PRIN ESTEEM and PRIN TENACE) and EU projects (EU-GUARDIAN, PANACEA, Semantic-Gov, ReSIST, GreenerBuildings, eDIANA, PANOPTESEC). Her main research interests are in the context of secure and dependable distributed systems. In particular, she targets topics like attack modeling, fault tolerance (also considering malicious faults) and management of autonomic process behaviors in unmanaged environment (e.g. peer-to-peer systems). Other research interests are: information dissemination systems, event-based systems, data dissemination in sensors networks.

Prof. Marco Angelini is an Associate Professor in Engineering in Computer Science at Link Campus University, Rome, and researcher at Sapienza University of Rome, Italy, Department of Computer, Control and Management Engineering, where he achieved his Ph.D. His main research interests include Visual analytics, applied in the Cybersecurity domain to provide situational awareness to cyber operators, and Cybersecurity, focused on designing solutions for cyber-defense of critical infrastructures against cyberattacks, security governance, assessment of cyber-risk and open-source intelligence. More about him can be found at: https://sites.google.com/dis.uniroma1.it/angelini