EconBPC: A Tool for Activity-based Monetary Assessment and Visualization of Security and Compliance Measures in Business Processes

Stephan Kuehnel [1], Stefan Sackmann [1], Johannes Damarowsky [1], and Martin Boehmer [1]

[1] *Martin Luther University Halle-Wittenberg, Universitaetsring 3, 06108 Halle (Saale), Germany*

Abstract

In this paper, we present the four-stage concept, implementation, application, and 2-stage evaluation of the tool EconBPC. EconBPC is a software artifact that emerged from a design science research initiative on the use of extensible event streams (XES). It was created to take a first step towards an automated activity-based monetary assessment of security and compliance measures in business processes and their addressee-oriented representation in a compliance view. For this purpose, EconBPC provides users with annotation features for XES and compliance processes, enables the storage of annotated log files, and the comparison of expenses for security and compliance activities with costs from associated regulatory breaches.

Keywords

Business Processes Compliance, Process Mining, Annotation, Cost, XES, event log, log file

1. Introduction

In the discipline of business informatics, business process compliance (BPC) addresses identifying, formalizing, implementing, checking, analyzing, and optimizing the adherence to business-relevant requirements before, during, or after executing business processes [1]. Due to the steadily increasing number of security and compliance requirements, e.g., the cross-sectoral provisions of the European Union's General Data Protection Regulation (EU GDPR), KRITIS for operators of critical infrastructures, or the second Payment Services Directive (PSD2) for banks, ensuring BPC has manifested itself not only as a complex technical challenge, but also as a cost-intensive task [2,3].

For example, current legislation on the protection of personal data requires the consideration of economic criteria when introducing compliance and security measures in business processes. Accordingly, Article 32 (1) EU GDPR defines that technical and organizational measures must be taken for the security of processing of personal data that ensure an adequate level of protection in relation to the respective risk and considering implementation costs. In contrast, a breach of Article 32 (1) EU GDPR may result in fines of up to 20 million euro or 4% of the total annual global turnover of the previous year (see Article 83 EU GDPR on "general conditions for imposing administrative fines"). In academia and especially in practice, the need for approaches and tools is communicated which no longer focus only on technical feasibility, but rather address economic aspects in ensuring BPC [4–6].

However, the economic assessment of BPC is a complex task for process owners, especially when data from large log files have to be analyzed. Studies from the field of decision making theory suggest the use of software artifacts to support such complex tasks in order to enable a reduction of the cognitive effort for the end user [7,8], e.g., by enabling special security and compliance views on business processes or by increasingly automating necessary calculations in large process models [6,9]. Currently, and to the best of our knowledge, neither are available, tools that explicitly visualize security and

Proceedings of the Best Dissertation Award, Doctoral Consortium, and Demonstration & Resources Track at BPM 2022, co-located with 20th International Conference on Business Process Management (BPM 2022), Muenster, Germany, September 11th to 16th, 2022.
EMAIL: stephan.kuehnel@wiwi.uni-halle.de
ORCID: 0000-0002-6959-9555

compliance measures of business processes in a separate view and tools that can economically assess and analyze these measures taking into account monetary cost and benefit aspects.

Consequently, the goal of our paper is to demonstrate the tool EconBPC, which, based on annotated extensible event streams (XES), takes a first step towards an automated monetary assessment of BPC and, moreover, can generate its own compliance view. The remainder of the paper is structured as follows. Chapter 2 shows the four-step concept of EconBPC, followed by the presentation of the software architecture and implementation in chapter 3. Chapter 4 illustrates a short application example and highlights the maturity of the artifact in a brief summary of the two-stage evaluation. The paper closes with some concluding remarks.

2. Concept of EconBPC

The economic principle is based at its core on an input-output relation that can be specified by quantitative parameters for a wide variety of domains, and thus also for BPC. The assessment of process-based security and compliance measures with regard to this relation, i.e. the economic assessment of security and compliance activities or activity sequences in business processes[2], therefore builds upon a graph-theoretical distinction between business activities and compliance activities. Consequently, we define a business process as a 3-tuple $G = (N, E, type)$, where $N = BA \cup CA \cup C$ represents a set of nodes in G that follow common execution semantics. $E \subseteq N\ x\ N$ represents a set of edges between nodes, i.e. a control flow, thus (N, E) constitutes a connected process graph. To simplify, we assume BA to be a set of purely business-related and CA to be a set of purely compliance-related activities. C represents a set of coordination nodes and the function $type$ assigns a coordination kind to each coordinating node of G, i.e.: $type : C \rightarrow \{start, end, split, synchronize, choice, merge\}$. In this sense, a compliance activity is not defined as an activity that itself complies with rules, such as the processing of customer data in a customer account, but as an activity that is carried out for the purpose of compliance, such as pseudonymisation of customer data in accordance with Article 4 (5) EU GDPR. Given that compliance activities have already been implemented in a business process to protect against a compliance risk and that the risk is associated with a specific damage value, compliance activities can be assessed economically. For this purpose, each compliance activity is specified in terms of costs and reliability (for a detailed description of the assessment approach and a calculation example, see [10]).

With the standard for XES, the IEEE Task Force on Process Mining has published an XML-based standard for log files that defines a generally accepted format for the exchange of log data between information systems and for the provision of data for analysis tools [11]. The XES standard can be extended by extensions and already offers a number of standard ones, such as the cost extension [12], but also specific ones, such as our extension "econbpc" (see https://bit.ly/2GI8U0E), which can distinguish activities into compliance-related and business-related, or specify activities by reliabilities.

The conceptual approach of EconBPC is shown in Figure 1 and includes four steps.

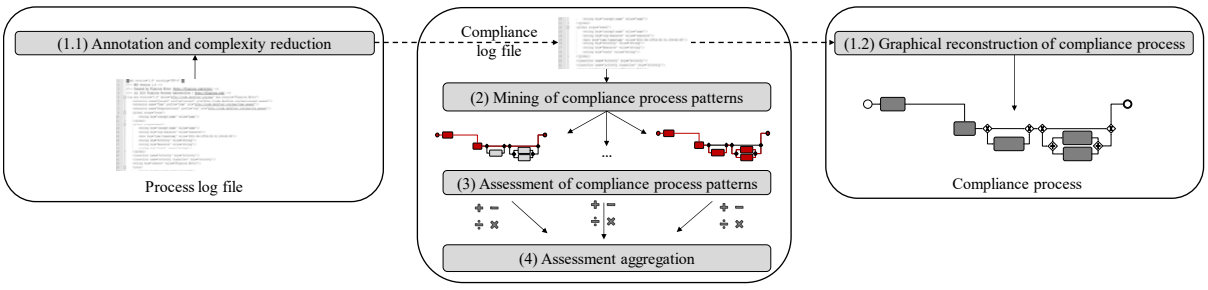

Figure 1: Conceptual approach of EconBPC

(1.1) Annotation and complexity reduction of log files. Given a process log file, we start with the first step. The flexible character of the XES standard opens up the possibility of annotating log files, i.e., enriching them with additional data [11]. This means, for example, that additional data can be stored in log files that further specify the activities performed (so-called events) [12]. If the available log file

[2] For ease of understanding, security and compliance measures are hereinafter summarized under the term compliance measures; the same applies to compliance activities, -processes, -patterns and -log files.

does not yet contain data about the event type (business/compliance event), costs, or reliabilities of compliance events, these must be annotated. This is followed by a complexity reduction. All events that are not of the "compliance" type are removed from the log file. The resulting so-called "compliance log file" thus only contains the data required for the monetary assessment of BPC, which increases the calculation efficiency of the procedure.

(1.2) Graphical reconstruction of compliance process. Although the approach does not require a graphical process reconstruction, the compliance process can be visualized based on the compliance log file. Such a visualization provides an overview of the compliance activities and can be used as a basis for presenting calculation results graphically in a way appropriate to the target audience, e.g., implemented via compliance activity annotations of costs or reliabilities for compliance officers.

(2) Mining of compliance process patterns. Each executed instance of a business process consists of a finite set of events and is referred to as a trace according to the XES standard [11]. A compliance log file contains only compliance traces, i.e. the executed instances of a compliance process. First, the event sequences of all compliance traces are analyzed and a list of all those sequences that differ in their sequence is created. Each entry in this list is a unique tuple that represents one of a finite number of pathways through a compliance process and is referred to as a compliance process pattern. For example, if two compliance traces do not differ, neither in the set of their events nor in their sequence, they are based on the same compliance process pattern. Second, the relative frequency of occurrence is determined for each pattern. If, for example, a compliance log file contains data from a total of three compliance traces and two of them do neither differ in their quantity of events nor in their sequence, then the relative frequency of occurrence of the underlying compliance process pattern is $66.\overline{6}\%$.

(3) Assessment of compliance process patterns. Using our computational approach originating from prior work [10], we determine the costs and reliabilities of the compliance process patterns[3]. Since compliance process patterns are derived from the traces of a business process and the control flow of a trace always has either a parallel or sequential character, only the computational rules for sequence, parallel split, and synchronization patterns need to be considered for assessment.

(4) Assessment aggregation. As part of the assessment aggregation, expected costs and expected reliabilities of the entire compliance process are calculated. For this purpose, we use the costs and reliabilities of the compliance process patterns and weight them with the frequencies of occurrence determined in step (2). At the highest level of aggregation, the expected costs of the compliance process can be compared with its benefits, which can be represented in terms of the monetary damage prevented, e.g., by avoiding regulatory breaches [10].

3. Software Architecture and Implementation

Design principles were used as a conceptual foundation for the implementation of the approach. Their derivation and the development of the software artifact build on the four steps presented in chapter 2. Design principles are understood as instructions or proposals for software artifacts that can serve to solve design problems and can be a conceptual precursor to artifact implementation [8]. Owing to space limitations, we refer to a previous study for the derivation of these principles (see: [13]).

Figure 2 shows the architecture of EconBPC[4] as a UML component diagram. The architecture model provides the basis for implementing EconBPC in R as a Shiny application. The import interface for log files enables the import of XES. The compliance data importer is an import interface for compliance data relevant to assessment, such as costs and reliabilities. The interface enables the annotation of log files with quantitative data of input and output factors of process-based compliance measures required for the calculation of monetary input-output ratios. The graph generator is used to reconstruct a business process graph from the log file (process discovery), which is output in the "process view" user interface. Based on an event filter, which is used to reduce the complexity of the annotated log file, the graph generator also enables the modular representation of compliance activities and activity sequences in the "compliance view" user interface. The path analyzer identifies the compliance process patterns,

[3] In the description of [10], the compliance process patterns are referred to as "paths". More details of the assessment approach have been omitted due to space restrictions. For a more detailed description, see [10].

[4] The R application EconBPC can be downloaded at the web address https://bit.ly/3OHN2Uh. A user guide and a screencast are available at https://bit.ly/2oXZtop and at https://bit.ly/2xwM2wW.

determines their relative frequencies of occurrence and sends a list of results to the assessment and analysis engine. Based on this, the engine enables: (1) the economic assessment of compliance process patterns and compliance processes, (2) the identification of inefficient compliance activities and compliance process patterns, and (3) the economic comparison of alternative compliance activities/processes. The aggregated valuation results are presented in the economic view user interface and can finally be used for decision support.

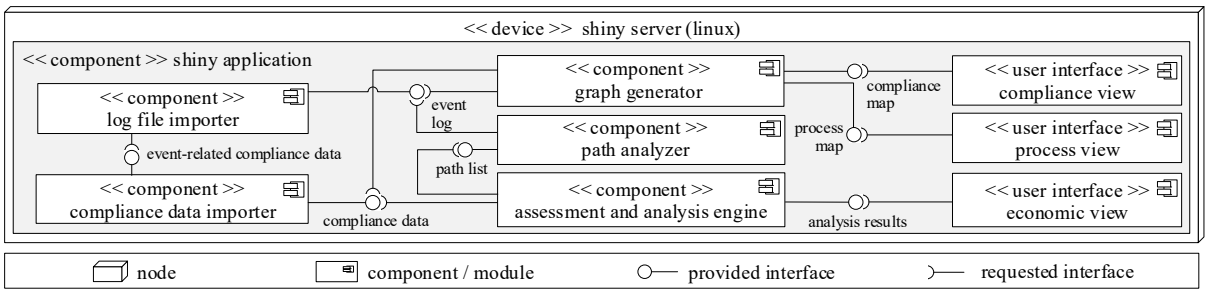

Figure 2: Software architecture of EconBPC represented as a UML component diagram

4. Application Example and Maturity

For demonstration purposes, EconBPC was tested with an artificial XES dataset publicly provided by Eindhoven University of Technology and included with the EconBPC download files. This is an anonymized dataset containing 105 synthetic instances of a credit application process. Since the dataset does not contain a specification of compliance events, the log file was annotated using the compliance data importer (see figure 3 (a)). "B1", "B2", "D", and "E" were specified as events of the compliance type and annotated with costs of $2000, $300, $5000, and $400, and reliabilities of 99%, 95%, 83%, and 100%, respectively. In addition, a fictitious compliance requirement was stored in EconBPC, the violation of which is accompanied by a fine of 30,000€.

(a) Screenshot of the log file/compliance data importer as well as the log file annotation feature

(b) Screenshot of the cost-annotated compliance process and calculation results

Figure 3: Screenshots of the EconBPC tool in the application example

Figure 3 (b) shows the compliance process of the application example in EconBPC as a multigraph, whose nodes have been annotated with cost data. Nodes where relatively high costs are annotated are highlighted in color to visualize potential inefficiencies. In the table below the compliance process, a list of the two identified compliance process patterns $(B1, D, B2, E)$ and $(B1, B2)$ can be seen (column "pattern"). The "absolute_frequency" and "relative_frequency" columns contain the absolute and relative frequencies of occurrence, and the "pattern_cost" and "pattern_rel" columns contain the costs and reliabilities of the compliance process patterns. The rows "expected cost", "expected reliability", "gross economic value", and "net economic value" contain the assessment aggregation results of the application example.

EconBPC has been evaluated twice. First, the design principles were evaluated by means of questionnaires in terms of perceived usefulness, comprehensibility, traceability and practicability (see

[13]). Second, the tool was tested by 12 experts from the field of IT governance, risk, and compliance in think-aloud sessions on case studies. The procedure of the think-aloud sessions followed a consistent pattern, initiated by 1) the provision of information about the software artifact and the tasks to be performed, 2) the systematic protocoling of the verbalized thoughts of the subjects, and 3) an analysis of the protocols. Positive expressions were marked with a "+", expressed potential improvements were marked with an "o", and each was assigned to the design principles. A discussion of both evaluations and related data can also be found in [13].

5. Closing Remarks

This paper has introduced the EconBPC tool, briefly highlighting its four-stage concept, its software architecture, its implementation as an R application, and its two-stage evaluation. By using XES, EconBPC offers the possibility to automatically visualize security and compliance measures of business processes in an addressee-oriented way and to assess them economically based on activities. Among many limitations stimulating future research, such as the subjectivity of the underlying conceptual and architectural design or the focus of the assessment on purely monetary values, great potential also lies in the transfer to real-world practice. As part of the project “Process-based Economic Evaluation and Selection of IT Security Measures” (ProBITS), which is funded by the Federal Ministry of Education and Research, we are already working on incorporating non-monetary values into the underlying calculation models, while still keeping the assessment simple enough to be scalable to different company sizes. However, testing in real case studies is still pending.

6. References

[1] E. Ramezani, D. Fahland, P. Mattheis, Separating Compliance Management and Business Process Management, Business Process Management Workshops, 2011, pp. 459–464.

[2] S. Sadiq, G. Governatori, Managing Regulatory Compliance in Business Processes, in: Handbook on Business Process Management 2, Springer Berlin Heidelberg, 2015, pp. 265–288.

[3] S. Sackmann, K. Kittel, Flexible Workflows and Compliance: A Solvable Contradiction?!, in: BPM - driving innovation in a digital world, Springer, Cham, 2015, pp. 247–258.

[4] J. Seo, K. Kim, M. Park, M. Park, K. Lee, An analysis of economic impact on IoT under GDPR, in: "ICT convergence technologies leading the fourth industrial revolution", Jeju, IEEE, Piscataway, NJ, 2017, pp. 879–881.

[5] A.M. Mustapha, O.T. Arogundade, S. Misra, R. Damasevicius, R. Maskeliunas, A systematic literature review on compliance requirements management of business processes, International Journal of System Assurance Engineering and Management 11, 2020, pp. 561–576.

[6] N. Adams, A. Augusto, M.J. Davern, M. La Rosa, On the Role of Process Mining in Business Process Compliance, SSRN Journal (2022). https://doi.org/10.2139/ssrn.4081558.

[7] Wang, Benbasat, Interactive Decision Aids for Consumer Decision Making in E-Commerce: The Influence of Perceived Strategy Restrictiveness, MIS Quarterly 33 (2009), pp. 293-320.

[8] H. Meth, B. Mueller, A. Maedche, Designing a Requirement Mining System, JAIS 16, 2015, pp. 799–837. https://doi.org/10.17705/1jais.00408.

[9] A.J. Varela-Vaca, L. Parody, R.M. Gasca, M.T. Gomez-Lopez, Automatic Verification and Diagnosis of Security Risk Assessments in Business Process Models, IEEE Access 7, 2019, pp. 26448–26465. https://doi.org/10.1109/ACCESS.2019.2901408.

[10] S. Kuehnel, A. Zasada, An Approach Toward the Economic Assessment of Business Process Compliance, in: Advances in Conceptual Modeling, ER 2018, Proceedings, Springer Cham, 2018, LNCS vol 11158. pp. 228–238.

[11] C.W. Günther, E. Verbeek, XES Standard Definition 2.0, second ed., Eindhoven, 2014.

[12] M.T. Wynn, W.Z. Low, t. A.H.M. Hofstede, W.E. Nauta, A framework for cost-aware process management: Journal of Universal Computer Science 20, 2014, pp. 406–430.

[13] S. Kühnel, S. Trang, S. Lindner, Conceptualization, Design, and Implementation of EconBPC – A Software Artifact for the Economic Analysis of Business Process Compliance, in: Conceptual Modeling, ER 2019, Proceedings, Springer Cham, 2019, LNCS vol 11788., pp. 378–386.