

Process Mining: Auditing Approach Based on Process Discovery Using Frequency Paths Concept

Zineb Lamghari*

Faculty of Sciences, LRIT associated unit to CNRST (URAC 29), Rabat IT Center Rabat, Morocco

In the company environment, the management team is responsible for producing normative models. The normative model is considered a standard model that aims at auditing all business processes in the same context. In this regard, the audit operation encompasses four process mining activities, in a hybrid evaluation (offline and online), which are the detect, the check, the compare, and the promote activities. This is still well performed for structured business processes. Otherwise, complex processes may deviate from the initial defined normative model context. Indeed, the latter must be refined for more precise results. Therefore, the combination of human knowledge, control-flow discovery algorithms, and process mining activities is required. To this end, we present a technique for reducing the complexity of unstructured process models (Spaghetti process models) into structured ones (Lasagna process models). This framework outputs a refined normative model for improving the future Business Process (BP) auditing operations. Moreover, this work introduces the sustainability advantage that can occur using process mining techniques.

Keywords: Process mining; Spaghetti process model; Lasagna process model; Normative process model; Control-flow Discovery algorithms

I. INTRODUCTION

Business Processes (BPs) are now an essential component of every organisational structure. They are set up to oversee and enhance the company's operations. In this context, information technology, such as Business Process Management (BPM) systems, Enterprise Resource Planning (ERP), Customer Relationship Management (CRM), and so on, aids in the automation of BPs (Nel & Abdullah, 2020). In this regard, information systems record BP execution-related event data in order to assess and steer issues involving the production of corporate value. Process Mining approaches were developed to attain these goals.

Process mining is a new scientific field that provides a link between computational intelligence and data mining. Also, it sits at link between process modelling analysis and computational intelligence. Process mining seeks to discover, monitor, and improve real-world processes by extracting knowledge from conveniently accessible event logs in

information systems. Indeed, the quality of event data is a crucial element in achieving Process Mining objectives. Process Mining typically assumes that execution data is stored in event logs. Indeed, event logs are the fuel for a Process Mining project and will be the difference between success and failure. In addition, as indicated in Figure 1, an event can be regarded as the primary step of Process Mining.

A process is made up of cases, or completed process instances. Each case is made up of a trace, which is a series of occurrences. Depending on the goals of an organisation, an event can include any number of extra qualities (timestamps, costs, resources, etc.). These extra characteristics, such as bottlenecks, which slow down the process flow, are crucial for various analyses (Corallo *et al.*, 2020). The forms of event logs might vary based on the information systems or aims. The quality of the event logs, on the other hand, is critical. The reason for this is that the outcome of Process Mining is highly influenced by the input (Van der Aalst, 2012).

*Corresponding author's e-mail: zineb_lamghari2@um5.ac.ma

Process Mining encompasses process discovery, conformance checking, and enhancement techniques (Van der Aalst, 2010). The process technique aims at generating a process model that mainly describes a business process based on such event data. The Conformance checking technique, which compares a pre-existing process model based on observable behaviour of event logs, to check whether reality conforms to the model. The Enhancement technique improves and extends an existing process model based on specific insights. Attempts to modify or expand an a priori model.

However, the conformance checking technique has been defined many times as an audit activity that evaluates a specific business process against the documented events and modules. In this context, auditing is a prominent operation aimed at verifying legal compliance in many sectors. In this sense, the refined Process Mining spectrum (Roubtsova & Wiersma, 2017) highlighted the audit activities in a hybrid context (online and offline), which are: detect (compares normative model with current events with the objective of detecting deviations at runtime); check (to identify deviations and measure the compliance level); compare (to assess how reality differs from what was intended or expected); and promote (upgrade parts of the revealed model to a refined normative model; current processes may be improved by boosting confirmed "best practices" to the de Initial Normative Model (INM)). Therefore, the audit operation is well done on Structured Business Processes (SBPs) based on the process discovery algorithms and the enterprise appropriate-normative model.

Otherwise, SBP and Unstructured BP (UBP) are the exact opposites of each other. Because of the structural intricacy of UBP, only a few process approaches may be used. These unstructured processes must be converted to structured ones to evaluate the process at hand and to promote parts of the discovered model. This can refine the normative model for quick auditing. By doing so, the combination of human knowledge, control-flow discovery algorithms, and audit activities of Process Mining is required. From the sustainability point of view, this helps in avoiding redundant and unnecessary work.

Indeed, simplification, promotion, and refinement actions are required, because organisations may struggle to respond

swiftly to day-to-day concerns if they do not have a comprehensive grasp of existing procedures and the capacity to adjust and monitor them. As a result, knowing where and how to enhance execution might be tough. This can require more computer resources and expose firms to a number of disadvantages, including the loss of a client and the possibility for repeat business, as well as greater employee turnover (Sonnenberg & Bannert, 2019). This is all relevant to sustainable development (Levina, 2015). Therefore, this paper's purpose is to provide a new approach for obtaining a refined normative model using process mining techniques related to sustainability meaning. Therefore, our paper addresses three major issues: 1) UBP's structural complexity is being reduced in order to make it more intelligible (converting UBP to SBP) in order to enrich the initial normative with new best practices. 2) Identifying the most frequently used execution paths such that the revised normative model may be obtained. 3) Describe how our approach works in tandem with sustainability.

The upcoming sections of this paper are organised as follows: Section 2 provides background and related work to our study field in terms of tools, techniques, and long-term benefits. Section 3 presents our proposed approach within a developed framework. Section 4 illustrates how it is possible to obtain a refined normative model by applying our framework steps. It also discusses the relationship between our approach and the sustainability goals. The conclusion is mentioned in section 5.

II. THEORETICAL BACKGROUND

In this section, we define the problems that are still faced related to the audit approach and the Process Mining field. They will be used throughout this paper. In addition, we will select a process mining tool and technique. Moreover, we will show in which sense our approach can support the sustainability standpoint.

A. Selection of the Audit Approach

Auditing occurs in a variety of circumstances (financial, maintenance engineering practices, health and safety issues, ethical conduct, etc.). We are focused on the context of process auditing, which is aimed at auditing a specific

business process against documented procedures. This can be solved by the Process Mining concept. In reality, Process Mining techniques are rarely used to support audits. The core technique is described in the improved Process Mining spectrum (Van der Aalst, 2016), where audit operations on structured BP may be performed. The most recent version of the auditing framework was updated in 2018 and employs an initial normative model to monitor and rectify the detected process model. By doing so, an audit report is obtained (which includes the final/refined normative model). None of the earlier publications revealed how to evolve from an initial normative model to a refined normative model by combining the initial one with the frequency paths concept. The frequency concept seeks to identify frequent execution paths within the process model (the discovered process model resulted from an UBP). It aids in path optimisation, performance optimisation, and resource management. Therefore, to achieve the audit progress related to Process Mining, we recommend combining the most recent auditing framework version (Roubtsova & Wiersma, 2017), where the initial normative model is considered as the human knowledge resulting from existing documentation, and the frequency paths concept.

B. Selection of the Process Discovery Algorithm

A process discovery algorithm constructs a generic process model based on event logs. Indeed, the generic model is an abstracted and general representation of real event logs. Several discovery algorithms are described with basic representations of process models, like the alpha algorithm. Other algorithms represent different abstraction levels, combined with clustering and classification techniques, to model processes from unstructured and complex events. In this sense, the authors conducted a comparative study to nominate the most preferable discovery algorithm. In this regard, the authors are impressed by the previous studies (Augusto *et al.*, 2018; Mans *et al.*, 2008; Pegoraro & Van der Aalst, 2019) to review the following process discovery algorithms: Alpha++, Heuristic Miner (HM), Inductive Miner (IM), Genetic Miner (GM), Fuzzy Miner (FM), State Based Regions (SBR), Language Based Regions (LBR), Language Based Regions (LBR), Language Based Regions (LBR).

On one hand, the Alpha ++ detects non-free choice relations by describing activities of the selected relation that depend on other activities (Wen *et al.*, 2007). It cannot detect invisible tasks. Therefore, this algorithm gives unsound results. In this sense, an extended version of the alpha algorithm has been created, to take into consideration the patterns' frequency. Indeed, the HM algorithm (Weijters & van der Aalst, 2003) can identify key actions and abstract exceptional and noisy ones while ignoring less important ones. This cannot be used to group traces that have a sub-log representation. Accordingly, the IM algorithm has been developed to treat events by grouping them into sub-logs. For each sub-log, a sub-process is generated. Then, a combination of the resulted sub-processes is released to obtain the generic process model. In this respect, the IM algorithm produces sound models (Bogarín *et al.*, 2018), i.e., fewer non-conformities are detected, and it fits with most present logs. However, it cannot identify complex and non-local process control patterns.

In this paper, we will use the Fuzzy miner algorithm (Günther & Van der Aalst, 2007), because it minimises the complexity found in a process model by emphasising significant information and dismissing less significant operations (fuzzy model). Fuzzy models may be transformed into many notations, such as BPMN, C-Net, and so on. For that purpose, we used fuzzy mining as the process discovery technique for creating a reduced process model representation. Therefore, the Fuzzy Miner algorithm gives sound results. On the other hand, the aforementioned algorithms have no initial and final markings, i.e., the model has no departure or arrival points, but it is readable. However, it helps understand links between different stages of the studied process.

On the other hand, new algorithms have been developed to treat event logs in their uncertainty, for example, the GM algorithm. This algorithm uses the genetic concept in creating process models from logs. This is done randomly. The precision metric is calculated for each process. Then, sound models are combined based on the mutation operation. The problem with this approach is its complexity in discovering and representing process models from real data sets (Vanden Broucke & Weerd, 2017). From the same complexity standpoint, FM deals with unstructured processes (Günther

& Van der Aalst, 2007). FM unstructured processes in this sense by preserving significant behaviour. While less important but highly linked behaviours are collected into clusters, less significant or less correlated activities are abstracted. The Fuzzy Miner algorithm is based on two concepts: significance, which measures the relative importance of each activity, and correlation, which measures the proximity of two successive activities. The significance of each activity can be evaluated according to its frequency, while the correlation can be defined by measuring the time of occurrence between two activities. Activities that occur shortly after each other are highly correlated. From these two concepts, it is possible to produce a simplified and coherent model using heuristics (Günther, 2009). The mined Fuzzy model contains primitive nodes, which contain only one activity, and cluster nodes, which contain activities. The Fuzzy Miner algorithm also offers users the possibility to zoom into or aggregate the model.

Furthermore, the SBR algorithm generates a Petri net from a Transition System (ST) based on specific abstractions, such as Set, Multi-Set, Sequence, and other types of abstractions, in which each state of the ST can be represented by a complete or partial trace. This algorithm ensures the fitness metric as well as the identification of complex control structures. Besides, SBR is unable to process incomplete and noisy logs (Van der Werf *et al.*, 2008), while the LBR algorithm can find process model places based on the language process. Indeed, the LBR algorithm uses properties derived from logs (causal relationships) to determine the final model by describing different places. Unfortunately, this algorithm is unable to process incomplete and noisy logs (Van der Aalst *et al.*, 2010). In summary, the primary purpose of a process discovery algorithm is to provide a high-quality model. There are several metrics and methodologies for estimating process model quality criteria (Günther & Van der Aalst, 2007; Buijs *et al.*, 2012), which are as follows:

- **Fitness:** This metric indicates how well the model captures the observed data. It measures how well a log fits into a model.
- **Generalisation:** The model should be able to interpret what is in the log. This metric evaluates the model's ability to explain unseen variables. The main difficulty

with the generalisation metric is the need for unobserved treatment.

- **Precision:** The model ignores unrelated data that is stored in the log. This metric estimates how much of the model has yet to be observed. A high level of generalisation model could represent much more than once presented in the log (underfitting model vs. overfitting model).
- The model should be represented in a simplified structure. This metric quantifies the model complexity, and it is not treated in this paper. For the fuzzy miner algorithm, the output model is a fuzzy model. To evaluate the fuzzy model, two metrics are available: node detail and conformance.
- The Node Detail describes how activities displayed in the Fuzzy model are related to the aggregated or deleted activities. Nodes of visible activities are called explicit nodes, while nodes corresponding to an activity are denoted as implicit ones.
- **Compliance** is a measure that describes the alignment between the fuzzy model F and the logs T . Each activity in the logs that does not exist in the Fuzzy model will be counted as a deviation.

C. Selection of the Process Mining Tool

In the trade, there are various commercial process mining tools, including Celonis (<http://www.celonis.de/>), Fluxicon (<http://www.fluxicon.com/>), and myInvenio (<https://www.my-invenio.com/>). However, PROM (Van Dongen *et al.*, 2005) is the most commonly used process mining tool in academia. Eindhoven University of Technology created PROM. It is an open-source framework that includes a number of process mining approaches. PROM provides an easy-to-use platform for corporate users as well as a framework for developers to create new extensions or plug-ins. PROM was chosen for using Process Mining techniques in this work because it provides a wide range of Process Mining approaches.

D. Sustainability Consideration

The most commonly used method to meet a sustainability goal in terms of reducing energy consumption was to re-think and re-design the general usage of resources (Thorsten *et al.*,

2017) such as electricity usage for lighting and cooling of the facility and the energy used for treating data. Based on this percent, to ensure a sustainable process, we must reduce processing errors: 1. Monitor processes for quick corrective reaction; 2. reduce computing resources when extracting the log information. Indeed, our approach passes through two stages. Stage (a) converts UBP to SBP. This reduces the complexity related to the treated data. As a result, the energy

consumed during the mining treatment process is influenced. Stage (b) optimises the resultant SBP to a refined normative model; this helps in monitoring business processes at runtime in the same context to instantiate corrective reactions. These two stages can be rented in terms of conserving energy, and this is the main objective of sustainability.

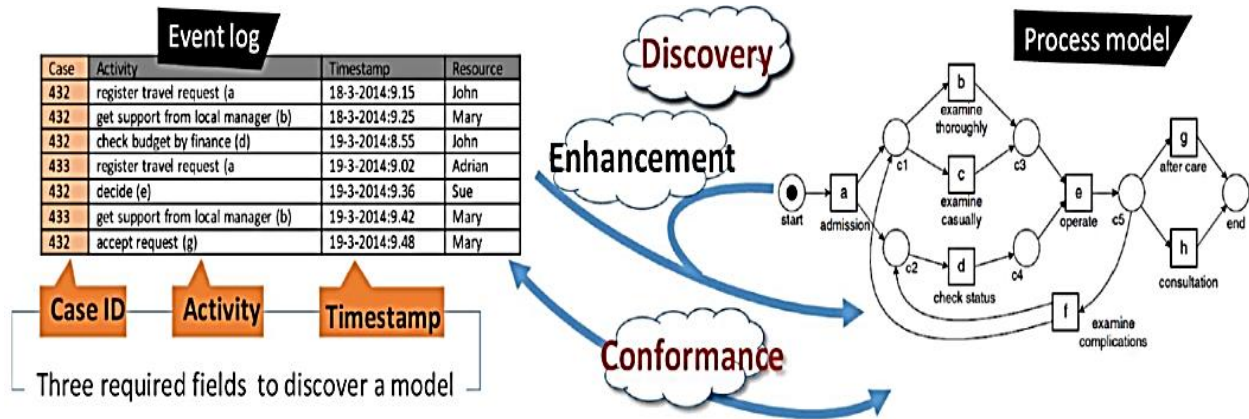


Figure 1. Figure caption must be typed with the normal style using “Times New Roman size 10 point”

III. FRAMEWORK

Our proposed framework for simplifying UBP to a refined normative model is given in Figure 2. Here, an event log of an UBP is read and simplified by applying a fuzzy mining technique or heuristic miner algorithm. From the fuzzy miner view, different variants of control-flow will be created with different percentages of activity and path percentage. This step is repeated until obtaining a concrete Lasagna structure with a fitness value at least equal to 0.8 (i.e., at least 80% of the behaviours should match between the model and the event log). This is included the check and compare activities (from the offline auditing point of view, we ignore the detect activity).

To refine the initial normative model, we will filter the obtained structured process model by the initial one (considered as human knowledge) based on (Roubtsova & Wiersma, 2017) for extracting control-flow related features. These features serve to identify the frequent execution paths with the suitable activities' abstraction based on the initial normative model.

IV. APPROACH APPLICATION

We utilise a live example from the road traffic fine collecting procedure to demonstrate the issues described in this article. Figure 3 depicts its control-flow process paradigm. It is made up of 11 activities, 56,140 occurrences, and 150 cases. The event log for traffic fine management is collected from the standard Process Mining repository (https://data.4tu.nl/articles/dataset/Road_Traffic_Fine_Management_Process/12683249).

We recommend using the fuzzy miner approach and the auditing approach to reduce the process structure from a spaghetti (Figure 3) to a lasagna process. In the first step, we use the fuzzy mining technique to simplify the UBP as a frequency matrix (Günther & Van der Aalst, 2007), after which we investigate deviations between the INM and the SBP and filter the SBP using the INM (Figure 4). This provides a suitable refined normative model for subsequent auditing operations. We defined the INM using these works as data sources (Van der Aalst *et al.*, 2015a; Mannhardt *et al.*, 2015b; Mannhardt *et al.*, 2017).

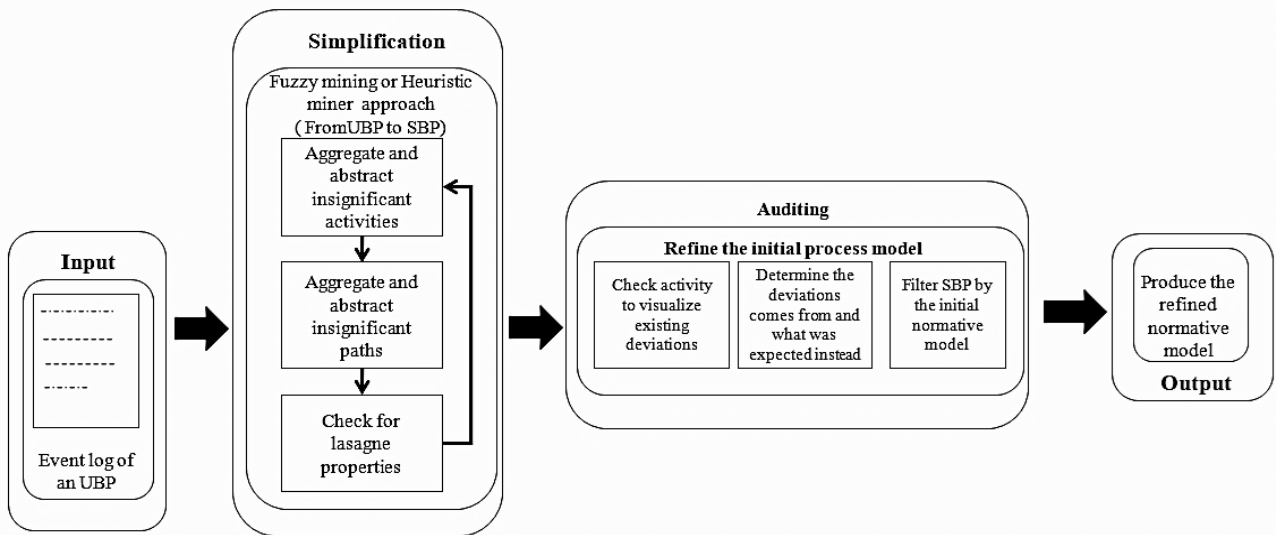


Figure 2. Standard Framework used for producing the refined normative model

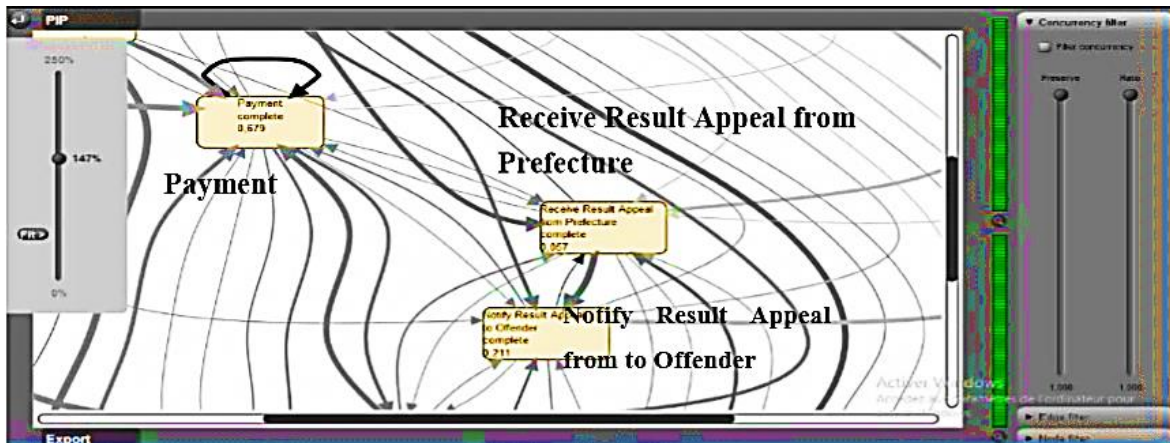


Figure 3. Spaghetti process: Excerpt showing 10% of road traffic fine management process

A. From UBP to SBP

In the first step, this paper adheres to the creator of Process Mining's idea (Van der Aalst, 2016). According to the idea (Van der Aalst, 2016), "a process is a lasagna process if it is feasible to develop an agreed-upon process model with a fitness of at least 0.8, i.e., more than 80% of the events happen as intended" with minimum effort. The road traffic fine management process is recreated using a fuzzy miner (Günther & Van der Aalst, 2007) at varying percentages of path and activity abstraction and aggregation. This process is continued until the premise's condition is met. It is obvious that decreasing the structural complexity of an activity and route enhances fitness value. The aim here is to discover the control flow with a fitness of at least 0.8. It is accomplished

by calibrating activity and path constituent percentages repeatedly during the procedure. Figure 4 shows a simplified Lasagna process that may be used to detect common execution pathways.

The Fuzzy algorithm is applied to our event logs using the Fuzzy Miner package. This algorithm shows that the visible nodes are important and significant (see Table 2). Moreover, replaying logs on obtained models gives sound results (conformance 0.8). Therefore, these models represent a clear and interesting direction. The frequency of users' choices is well presented and provides a useful overview of the studied process.

B. Check and Compare

At this stage, we investigate where and why the INM deviates from historical occurrences using the Prom tool's conformance checking package. Then, we compare this initial normative model to the SBP to determine these deviations. This can help to release a promotion. This shows discrepancies between the INM and the de facto model (see

Figure 4): 1-Insert the date of the prefecture's appeal; 2-Send for credit; 3-Send the appeal to the prefecture; 4-Receive the prefecture's decision; 5-Notify the offender, and 6-Appeal to the judge. Finally, the INM will be advanced by using the SBP model to create a revised normative model. For example, new activities are added in comparison to the INM to achieve the extra penalty activity.

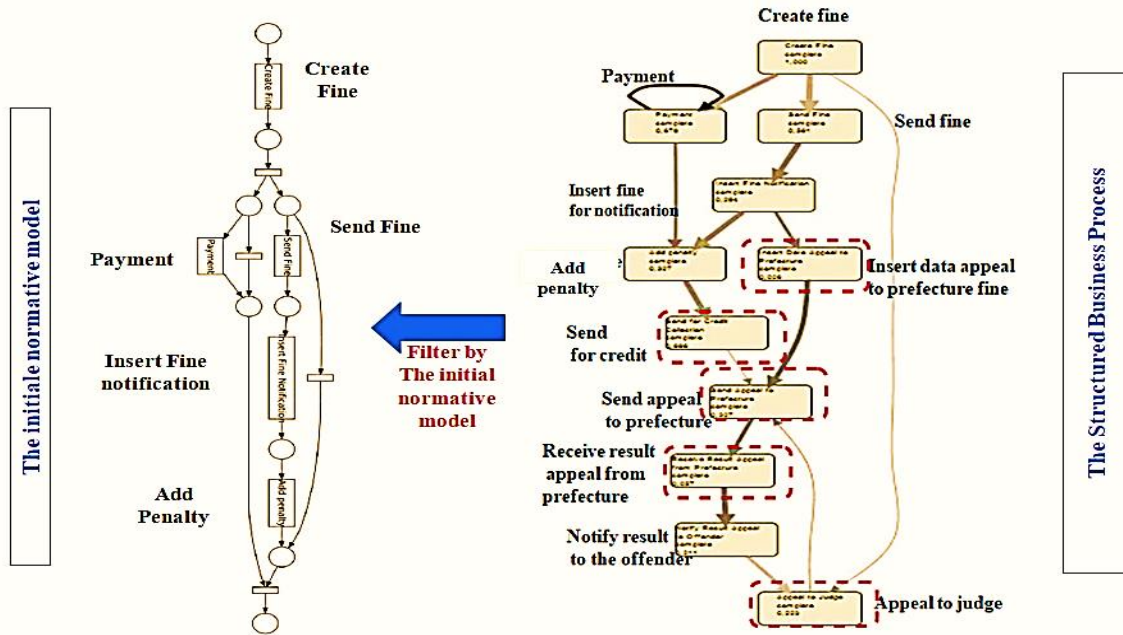


Figure 4. Filtering the SBP by the initial normative model

Table 1. Frequency matrix based on the initial process model filtering (x followed by y)

x>y	CF*	SF*	I*	A*	P*	SP*	ID*	RP*	NO*	AJ*	SC*
CF*	0	1	0	0	1	0	0	0	0	0	0
SF*	0	0	1	0	0	0	0	0	0	0	0
I*	0	0	0	1	1	1	0	0	0	0	0
A*	0	0	0	0	1	0	0	0	0	0	1
P*	0	0	0	0	0	0	0	0	0	0	0
SP*	0	0	0	0	0	0	1	0	0	0	0
ID*	0	0	0	0	0	0	0	1	0	0	0
RP*	0	0	0	0	0	0	0	0	1	0	0
NO*	0	0	0	0	0	0	0	0	0	1	0
AJ*	0	0	0	0	1	0	0	0	0	0	1
SC*	0	0	0	0	0	0	0	0	0	0	0

C. Promote

In this phase, the SBP model will encourage the INM throughout this step to produce the final/refined normative model. This improved model will be the new standard auditing process model.

Table 2. Measures for the fuzzy miner algorithm

	Node details	Conformance
Road traffic event logs	1.00 (equal to 100%)	0.80 (EQUAL TO 80%)

The frequency matrix (showing the percentage of direct flows between activities) is fully incorporated into the fuzzy miner algorithm application. Furthermore, differences between the INM and the SBP are found by comparing and verifying activities, as a result, we go straight to the refining action, filtering the SBP with the INM into the Promote activity. This results in a new frequency matrix from which an improved normative model may be generated (see Figure 5). In this sense, we define relationships between the finding activities by using two values (1 and 0) which are (solid or not).

The activities depicted in Figure 3 are the outcomes of the filtering phase (promote based on observed deviations), i.e., the full log activity called CF has a strong association with the SF activity. For example, CF as x is immediately followed by SF as y ($x > y$).

According to Table 2, new activities are observed and must be handled by the refined normative model (Cf. Figure 5), which includes: CF (Create Fine), SF (Send Fine), I (Insert Fine Notification), A (Add Penalty), P (Payment), SP (Send Appeal to Prefecture), ID (Insert Date Appeal to Prefecture), RP (Receive Result from Prefecture), NO (Notify Offender), AJ (Appeal to Judge). As a result, the revised process involves the following steps: If the offender files an appeal within 60 days of receiving the fine, the appeal is submitted to the associated prefecture (Send Appeal to Prefecture) and is registered when it is received (Insert Date Appeal to Prefecture). The outcomes of the appeal are returned to the municipality (Receive Result Dispute from Prefecture) and communicated to the offender (Notify Result Appeal to Offender), who has the right to appeal the result

(Appeal to Judge). If the offender fails to pay (perhaps because of a refused appeal), the fine is transferred to credit collection (Sent for Credit Collection).

Petri Net notation is used to depict the enhanced normative model. This notation is appropriate for the following verification. Figure 5 depicts a model that does not represent any type of parallelism. It is just a series of tasks interspersed with options. This notation is appropriate for the following verification. Figure 5 depicts a model that does not represent any type of parallelism. It is just a series of tasks interspersed with options.

To summarise, the input data (event log) was arranged to facilitate the discovery effort by categorising the log into deviated and non-deviated activities to generate separate sub-logs. These sub-logs are needed to carry out the workflow's actions.

As a result, the simplicity and structure of the resulting road traffic process illustrate the efficiency of the transfer from UBP to SBP. The control-flow was sophisticated and had a complex structure, as seen in Figure 3. After the transformation, it appears in a legible form with a decreased number of activities and pathways (see Figure 5).

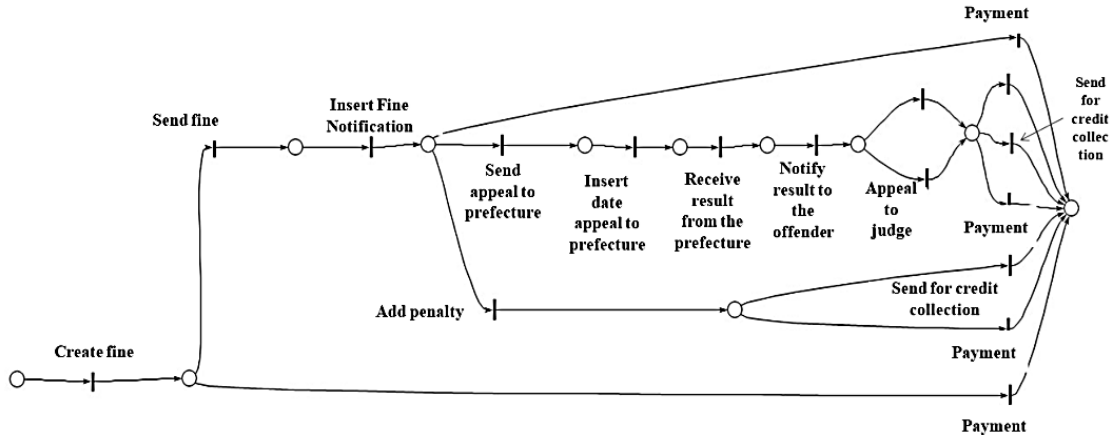


Figure 5. The refined normative model designed in Petri Net

D. Sustainability Point of View

Generally, our approach has provided a refined normative model, which aims to achieve the auditing operation. This is done by detecting violations in order to instantiate a corrective reaction. This is accelerating the process execution.

In this sense, it is clear that the sustainability concept has been defined into three main pillars, which are: the tool used, the resource consumption, and the technique followed, where the monitoring technique can define the type and quantity of data treated into a process mining tool, and this gives insights on the energy consumed by a resource. So, the data complexities, which are combined with a suitable tool, could influence resource consumption and, by default the sustainability concept. This is possible and could be rentable if we combined this approach with the three main pillars of sustainable development (Social, Economic and environment) in an auditing context. So, our process mining approach has a positive impact in terms of sustainability. To achieve a positive balance, our approach aims to avoid redundant and unnecessary work (human), communication (network) and computing resources when extracting the log information. Therefore, a trade-off between extracting only the data that is needed and extracting data in a way that minimises manual setup work, should be found.

V. CONCLUSION

In this paper, we have presented an auditing approach within a developed framework of different steps. Our approach aims at generating a refined normative model for the auditing of

BP treatment. This is done by passing from an unstructured BP to a structured BP using the fuzzy mining algorithm. Then, we filtered the resultant SBP by the initial normative process model using the frequency paths concept. These two strategies have been proven by testing them on real-world event logs. The refined process model helps in reducing the computing energy consumed during BP auditing treatment by simplifying unstructured business processes and reducing computed event log complexity. This is the advantage of our approach in terms of sustainability.

As further work, we plan to develop an executive plug-in in terms of an automated discovery algorithm based on the Augusto *et al.* approach (Augusto *et al.*, 2018). This will automatically structure each resultant UBP.

VI. ACKNOWLEDGEMENT

This work is supported by the National Center for Scientific and Technical Re-search (CNRST) in Rabat, Morocco.

VII. REFERENCES

- Augusto A, Conforti R, Dumas M, La Rosa M & Bruno G 2018, 'Automated discovery of structured process models from event logs: The discover-and-structure approach', *Data & Knowledge Engineering*, vol. 117, pp. 373-392.
- Bogarín Vega A, Cerezo Menéndez R & Romero C 2018, 'Discovering learning processes using inductive miner: A case study with learning management systems (LMSs)', *Psicothema*, vol. 30, no. 3, pp. 322-329.
- Buijs JC, Van Dongen BF & van Der Aalst, WM 2012, 'On the role of fitness, precision, generalization, and simplicity in process discovery', in *On the Move to Meaningful Internet Systems Confederated International Conferences*, Springer, Berlin, Heidelberg.
- Corallo A, Lazoi M & Striani F 2020, 'Process mining and industrial applications: A systematic literature review', *Knowledge and Process Management*, vol. 27, no. 3, pp. 225-233.
- Li J, Liu D & Yang B 2017, 'Process Mining: Extending α -Algorithm to Mine Duplicate Tasks in Process Logs', *Advances in Web and Network Technologies, and Information Management*, pp. 396-407.
- El-Gharib NM & Amyot D 2019, 'Process Mining for Cloud-Based Applications: A Systematic Literature Review', in the *IEEE 27th International Requirements Engineering Conference Workshops*.
- Günther CW & Van der Aalst WMP 2007, 'Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics', *Business Process Management*, pp. 328-343.
- Levina O 2015, 'Exploring the Role of Business Process Management in Sustainability Initiatives', *Mediterranean Conference on Information Systems*.
- Mannhardt F, de Leoni M, Reijers HA & van der Aalst WMP 2015a, 'Balanced multi-perspective checking of process conformance', *Computing*, vol. 98, no. 4, pp. 407-437.
- Mannhardt F, de Leoni M & Reijers HA 2015, 'Extending Process Logs with Events from Supplementary Sources', *Lecture Notes in Business Information Processing*, pp. 235-247.
- Mans RS, Schonenberg MH, Song M, Van der Aalst WM & Bakker PJ 2008, 'Application of process mining in healthcare—a case study in a dutch hospital', in *International Joint Conference on Biomedical Engineering Systems and Technologies*, Springer.
- Martin, N *et al.* 2020, 'Recommendations for enhancing the usability and understandability of process mining in healthcare', *Artificial Intelligence in Medicine*, vol. 109, series. 101962.
- Nel P & Abdullah MA 2017, 'Mediating effects of business success on the relation between business issues and business prospects amongst immigrant entrepreneurs in Malaysia', *Contemporary Management Research*, vol. 13, no. 1, pp. 3-18.
- Pegoraro M & van der Aalst W M 2019, 'Mining uncertain event data in process mining', in *2019 International Conference on Process Mining (ICPM)*, IEEE.
- Roubtsova E & Wiersma N 2018, 'A Practical Extension of Frameworks for Auditing with Process Mining', *Proceedings of the 13th International Conference on Evaluation of Novel Approaches to Software Engineering*.
- Thorsten S, Behrens D & Knackstedt R 2017, 'Sustainability in Business Process Models: A Taxonomy-Driven Approach to Synthesize Knowledge and Structure the Field', *Proc. of the Int. Conf. on Information Systems*.
- Van Dongen BF, de Medeiros A, Verbeek H, Weijters A & van der Aalst WMP 2005, 'The ProM Framework: A New Era in Process Mining Tool Support', *Lecture Notes in Computer Science*.
- Van der Aalst WMP 2010, *Process mining discovery, conformance, and enhancement of business processes*, 1st edn, Springer.
- Van der Aalst WMP, Daniel F, Barkaoui, K & Dustdar S 2012, 'Process Mining Manifesto', *Business Process Management Workshops*, vol. 99, pp. 169-194, Berlin, Heidelberg: Springer, Berlin, Heidelberg.
- Van der Aalst WMP 2016, 'Operational support in Process Mining: Data Science in action', pp. 301-317, 2nd eds, Springer, NY.
- Van der Aalst WMP, Bolt A & van Zelst SJ 2017, 'RapidProM: mine your processes and not just your data', preprint of 2017.

- Vanden Broucke SK, & De Weerd J 2017, 'Fodina: A robust and flexible heuristic process discovery technique', Decision support systems, vol. 100, pp. 109-118.
- Van der Werf JME, Van Dongen BF, Hurkens CA & Serebrenik, A 2008, 'Process discovery using integer linear programming', in International Conference on Applications and Theory of Petri Nets, Springer.
- Weijters A & van der Aalst WMP 2013, 'Rediscovering Workflow Models from Event-Based Data using Little Thumb', Integrated Computer - Aided Engineering, vol. 10, no. 2, pp. 151-162.
- Wen L, Van Der Aalst W M, Wang J & Sun J 2007, 'Mining process models with non-free-choice constructs', Data Mining and Knowledge Discovery, vol. 15, no. 2, pp. 145-180.