Journal of Al-Farabi for Engineering Sciences (JFES)
Vol. 5, No. 1, 2026, **pp**. 24-41
ISSN: 2617-5517 | **DOI**: https://doi.org/10.59746/jfes.vxx.xx

Process Mining and Business Intelligence for the Detection of Suspicious Transactions

Fatimtzahraa Ghassan Majid[1], Huda Lafta Majeed[2,*], Gromov, Y.Y[3]

[1,2] Computer Science and Information Technology, University of Wasit, Iraq. Email:stdm.fghassan@uowasit.edu.iq ; hulafta@uowasit.edu.iq
[3]Director of the Institute of Automation and information technology, Tambov State technical university, Russia. Email: gromovtambov@yandex.ru

Received 22/01/2026, Revised 18/02/2026, Accepted 21/03/2026

Abstract Digital forensic investigations still have a hard time finding suspicious activity in huge financial datasets. When it comes to finding complicated behavioral patterns and little abnormalities in big transactional databases, traditional database forensic methods have their limits. This is because they typically depend on human analysis or searches that have already been set up. This work presents a data forensic system grounded on process mining to assist in identifying suspicious activity inside financial systems. The framework has three main parts: process discovery, which uses Petri net representations to make diagrams of processes to event logs; conformance checking, which at first compares the data observed with the uncovered method framework; and clustering and behavior analysis, which puts suspicious cases at the top of the list based on unusual traces. We evaluated the suggested method on an actual-world financial database and put it through its paces utilizing the ProM mining framework. The trials showed that the strategy works to find strange patterns of transactions while making it easier to find any fraudulent activity. The suggested process mining-based structure is better than existing methods since it is organized, scalable, and automated.

Keywords: Process Mining, Database Forensics, Anomaly Detection, Petri Nets, Digital Forensic Analysis, Business Intelligence, Digital Database Forensics.

1. Introduction

Due to its rapid expansion and revolutionary impact on financial services, digital banking has revolutionized the way money is transacted online. Security issues have also intensified due to the shift to digital. Banks and various other financial businesses are increasingly vulnerable to cyber risks such as financial fraud, unauthorized access, and identity theft. Finding questionable transactions and investigating digital evidence has therefore become critical for ensuring the security of financial systems.

Digital forensics is particularly crucial for identifying, collecting, and judging digital proof of crimes. databases forensics has grown into an important discipline in this domain since massive transactional databases are such a significant part of modern financial systems. Investigators usually look at transaction data, access patterns and system events to uncover suspected fraud. [1][2] But traditional database forensic approaches mainly rely on people looking at things, static queries, and rules which have already been created. These approaches take a long time, are easy to mess up, and are not always successful when you need to uncover little problems or sophisticated patterns of behavior in large, changing datasets. [3][4]

Recent research has examined the use of process-based technology for data analysis to mitigate these constraints. To uncover genuine workflows in event logs, process mining uses data mining and modeling of systems to look at them. [5] Process mining shows how corporate processes really work, which lets investigators see how transactions flow, find discrepancies, and spot abnormal

behavior patterns. [6][7] Process mining looks at full execution traces instead than just single events, which gives a better picture of how a system works than older forensic methods. [8][9] [10]

Using these skills, this research suggests a procedure based on mining forensic approach for discovering questionable transactions inside financial datasets. Anomaly assessment, clustering, process discovery, and compliance validation are all parts of the suggested method that may be used to find and rank suspicious patterns of activity. To show that the framework is useful for modern banking criminal investigations, it is tested with actual data on financial transactions. These sources: [11] [12].

Lots of money changes hands, and those changes are only becoming more intricate. Conventional approaches to database forensics struggle to keep pace. They mostly depend on manual evaluation or static searches, which aren't up to the task of spotting fraud or complicated patterns in big, dynamic databases. This highlights the need for a real-time analytics-based automated system to model and assess financial processes. Therefore, more intelligent and adaptable technology is needed for digital forensic investigations into the financial system to be successful.

Despite the effectiveness of process mining strategies in enhancing processes in several sectors, database forensics is still not sufficient research emphasis. This study tries to rectify that by applying these methods to this domain. Through the integration of process identification, compliance identification, behavioral clustering techniques, and rarity assessment, this study provides a forensic tool for financial information that is both comprehensive and user-friendly. Using this model might be helpful in ranking the seriousness of odd dangers. The method's efficacy and widespread applicability have been shown via testing on actual financial data. It aids in automating and improving fraud detection as well.

This study makes several contributions to the field of digital database forensics:

1. It proposes a structured forensic framework that integrates multiple process mining techniques to analyze transactional behavior in financial databases.
2. It combines process discovery, conformance checking, and clustering methods to systematically identify deviations from normal transaction flows.
3. It introduces a rarity-based prioritization mechanism that ranks anomalous traces according to their potential forensic relevance.
4. It demonstrates the effectiveness of the proposed framework through experimental evaluation using real-world financial data and the ProM process mining environment.

2. Related Work:

According to studies, digital and databases forensics are becoming more crucial in online banking as more and more financial institutions switch from paper-based to digital systems. Financial institutions are now more vulnerable to cyber threats such identity theft, data manipulation, and illicit access, despite the fact that this development has made transactions simpler to access [13].

Digital forensics has created structured ways to collect and analyze digital evidence. It has become an essential tool for aiding the courts and police. Database forensics is very important in this situation since financial transaction records kept in databases are very sensitive and need to be protected against tampering or abuse. The size, spread, and constant change of today's databases, on the other hand, make forensic methods harder to use [14] [15].
Database forensics has to cope with a number of technical and organizational issues. For instance, it could be impossible to get proof without messing with operating systems, attackers generally attempt to cover their tracks, and encryption and memory architectures that can be changed make things much tougher. The absence of standardized methodologies and instruments for research

databases diminishes the reliability of results and reduces their likelihood of acceptance in legal proceedings [16] [17].

Today, forensic approaches rely on things such as log analysis, periodic photos, triggers, or continuing analysis. Each has clear restrictions, but they are useful. For instance, they depend on complete data, lack a temporal dimension, impede system performance, and fail to represent complex behavioral patterns in extensive financial systems [18].

In this situation, process mining is a new way to analyze data that leverages event logs to integrate data mining with modeling processes. With this method, you may find real transactions, check for compliance, and do analysis in context. It has been useful in banking for finding fraud and mistakes in procedures that traditional forensic methods could miss [19] [20].

Two examples of conventional analytical frameworks used to examine databases are integrity evaluation and probabilistic approaches (Bayes and Dempster-Shafer). But they generally work on their own and don't reflect how processes change over time to be a whole [21]. Research reviews validate the deficiencies of existing tools regarding scalability, integrating, and automation, highlighting a significant gap in research focused on the incorporation of process mining techniques into a cohesive database assessment framework capable of handling extensive and varied records while proficiently identifying anomalous behavior [22].

Digital and databases forensics has come a long way, but there are still certain problems with the methods that are now available. Conventional forensic methods rely mostly on static query and human involvement; they include log analysis, photographs, trigger-based surveillance, and manual inspection. These approaches aren't up to the task of dealing with the massive amounts of complicated financial transaction data that are available today, and they can't even begin to scratch the surface of dynamic datasets for signs of fraud or unusual activity.

A comprehensive picture of process evolution over time is also lacking from many analytical methods, including probabilistic models and conventional data mining tools, which function independently. Despite process mining's widespread success in fields like healthcare and BPM, it has only recently found a home in database forensic investigations [23].

Therefore, a clear research gap exists in developing an integrated, automated, and scalable forensic framework that combines process mining techniques with database forensic analysis to model transactional behavior, detect anomalies, and prioritize suspicious activities effectively in large-scale financial systems.

2.1 Digital and Database Forensics Studies

Cybercrime investigations and digital evidence collection for legal and safety investigation need digital forensics. Database forensics analyzes databases, activity logs, and data to recreate events and identify illicit activity [24]. stressed metadata contexts in data investigations and showed how forensic analysis might help recreate database evidence [2]. analyzed many database forensic technologies and discussed the difficulties of obtaining meaningful evidence from big and complicated databases [25] [3].

Analysis of transaction logs, database snapshots, trigger-based surveillance, or live system analysis are some of the classic methods used in database forensics. Investigations may be reconstructed and suspicious alterations identified using these procedures. The approaches in question rely on human inspection, can't scale well, and struggle to manage remote and high-volume financial datasets, according to earlier research [1][26][27]. Additionally, database forensic investigations

for real-world financial systems are not as efficient or reliable due to the absence of standard forensic framework and automated analytical tools [28][29].

2.2 Process Mining Studies

Process mining is an emerging analytical technique that integrates data mining with process modeling by extracting knowledge from event logs. The concept was extensively developed by van der Aalst and colleagues, who demonstrated how event logs can be used to discover, monitor, and improve business processes [4][30][31]. Process mining techniques have been widely applied in several domains such as healthcare, education, manufacturing, and business process management to analyze operational behavior and detect process inefficiencies or deviations [5][6] [16].

Several algorithms have been proposed for process discovery, including Alpha Miner, Heuristics Miner, Inductive Miner, and ILP Miner. These algorithms generate formal process models, often represented using Petri nets, which enable the analysis of sequential and concurrent process activities [10][29]. Conformance checking techniques can then compare the discovered process model with actual event logs to detect deviations or inconsistencies in process execution [12]. Although these techniques provide powerful capabilities for analyzing complex event-driven systems, their application in digital forensic investigations—especially database forensic analysis—remains relatively limited [15].

2.3 Fraud and Anomaly Detection Studies

There has been a lot of research on finding fraud and other strange behavior in the fields of cybersecurity and banking and finance. A number of research employ statistical structure, machine learning approaches, or pattern mining methods to find strange transactions and suspicious activity. For instance, look into frequent pattern mining methods. They look for trends in normal transactions and point out any differences that could point to fraud [31][32] [33]. There are various ways to use probabilistic analysis models, such Bayesian or Dempster-Shafer theory, to figure out how likely it is that suspicious database occurrences will happen [23].

These approaches work well for discovering outliers in information about transactions, but they typically see transactions as random events instead as parts of a bigger picture. As a result, people might not be capable to understand the complex behavioral links that exist in banking systems between activities that happen one after the other. Moreover, some anomaly detection methodologies just identify atypical events, rather than facilitating structured forensic investigations or prioritizing dubious situations [30][34].

2.4 Research Gap

Although there has been a lot of advancement in digital forensics, processing mining, and fraud detection methods, the majority of the current research still treats these topics separately instead of combining them into one cohesive analytical framework. Complex patterns of behavior in large-scale financial information databases might be difficult for traditional database forensic methodologies to discover due to their reliance on static query analysis and human examination. Simultaneously, methods for detecting fraud often examine out-of-the-ordinary transactions in isolation from the larger systemic context.

Integration of process mining with database forensic analyses is still restricted, despite process mining's excellent capabilities for modeling event-driven systems and discovering deviations via conformance analysis. Forensic analysis in financial systems is a growing field, but there has been little research on scalable frameworks that integrate process discovery, compliance verification, behavioral grouping, and anomaly prioritizing.

As a result, there has been little effort to create a scalable framework for investigating banking system databases that combines process discovery, compliance verification, behavioral grouping, and anomaly prioritizing, despite advances in process mining and digital forensics.

3. Proposed Approach

This research introduces a controlled automatically data-driven forensic approach using procedural mining tools to investigate suspicious patterns inside database systems, particularly in banking environments. The concept tries to fix the problems with current database forensic methods, which include manual analysis, limited scaling, and subjective interpretation, by combining sequences of events concerning database records and carefully looking for any discrepancies. Using method identification, conformance confirmation categorization, and advanced deviance analysis [25], the technique helps investigators locate, rank, and analyze unanticipated behaviors more quickly and accurately.

The proposed technique [26] is based on the Fraud Evaluation approach. A common approach to discover fraud is this. This technique of looking at information from events makes it easier for the framework to discover evidence of fraud. This framework automatically and fully encompasses features that are related to time, behavior, and organizations. This is distinct from the norm in that it looks at beyond a single event or a small number of points of data. This helps investigators identify hidden patterns and alterations to operations that might hurt intricate financial systems.

That Framework and Basic Structure for Event Logs The proposed method depends on three aspects that work together to get the most out of it:

1. Run a simulation that produces genuine process flows from event logs;
2. clustering, which finds differences by grouping tracks that are similar and then
3. In order to discover discrepancies and determine what to do when something appears odd, analysis makes use of figures and graphs. Together, they improve the whole forensic process, which speeds up the inquiry and makes it easier to find items.

The Proposed Approach part has been thoroughly rewritten to enhance its scientific clarity and rigorous academic content. We used exact scientific language to rewrite vague or informal statements. The wrong phrases "Petri net" and "Alpha Data Mining" were changed to the right ones: "Petri net" and "Alpha Miner algorithm." Also, the numerical definitions were finished, and the necessary formal written language was added to make sure that the framework description was consistent. The technique is now described in a simpler step-by-step way, going over the key parts of the suggested system, such as preprocessing the event log, finding processes, testing for compliance, clustering behavior, and prioritizing anomalies. These changes make the proposed forensic system easier to understand, more accurate, and easier to reproduce. They also make the section more eligible for publishing.

After being collected from audit, transaction, and system log files, data is preprocessed to rectify errors, standardize timestamps, fill in missing values, and link events to their appropriate contexts. This is all done to make the data more accurate and consistent. After that, a model similar to a Petri dish is built by taking the processes out of the data using methods like the Inductive Miner and Alpha Data Mining [27][35]. Petri networks are effective for behavioral evaluation that helps with tracking and validation since they have formal semantics and can show control flow [28] [29]. Figure 1 demonstrates how Petri networks may be used to depict actions and dependencies by using the example of booking a hotel.

 ISSN: 2617-5517 | **DOI**: https://doi.org/10.59746/jfes.vxx.xx

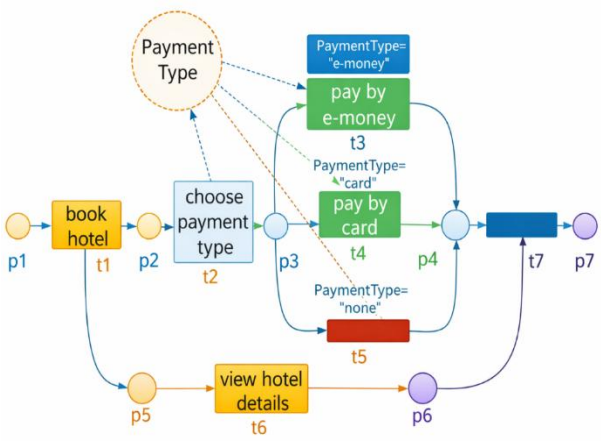

figure 1: Decomposing Petri nets for Book Hotel Process Mining

Conformity verification checks to determine whether the movement in the event log corresponds to what the Petri network says it should be once the model is generated. Metrics like fitness, precision, and violations are used to verify for alignment [30] [31][32]. This step sets up two logs: a corresponding log that indicates normal behavior and an outlier log that reveals difficulties and deviations. The following phase in the forensic inquiry is based on these data (Figure 2).

Figure 2 depicts the whole mining processes methodology that was utilized in this paper. First, the event log is filtered and preprocessed. Then, process recognition and compliance diagnostics are done. This illustration depicts how people meticulously clean up and look over raw database logs to determine the difference between operations that are in line with the rules and those that are not. This is what we will use to organize and analyze the data.

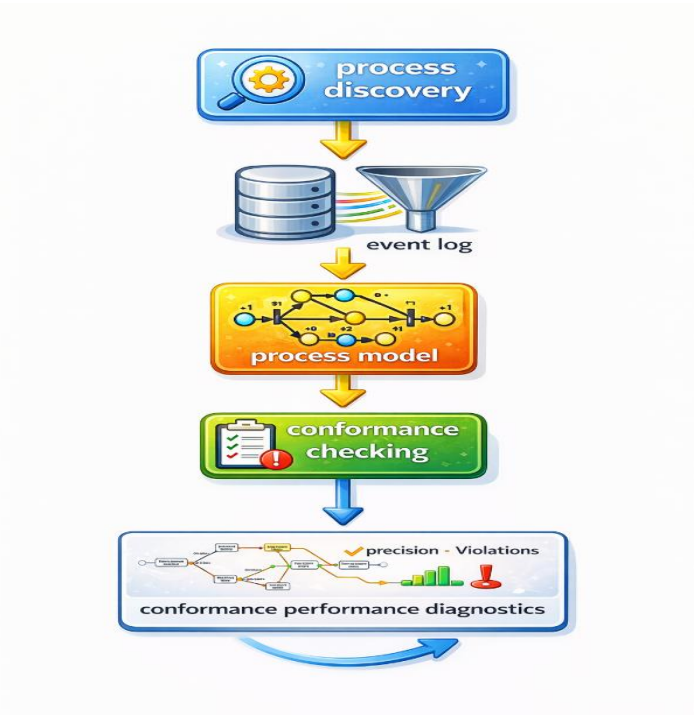

Figure 2: Process Mining Workflow: From Event Log Filtering to Conformance Diagnostics.

3.1 Clustering Component

The clustering component focuses on organizing traces based on behavioral deviations from the discovered process model. Although not every deviation implies fraud, deviations are treated as forensic indicators that warrant further examination [33]. Based on the nature of deviations, traces are categorized into four types:

- Skipped Trace (ST): one or more expected activities are missing;
- False Pattern Trace (FPT): activities occur in an incorrect order;
- False Decision Trace (FDT): an incorrect decision outcome is reached;
- False Throughput Trace (FTT): activity duration deviates significantly from expected values [34].

Formally, let Abe the set of activities in a trace σ, where each activity has a start time $t_s$and completion time t_c. A skipped trace is identified if an expected activity is missing:

$$ST(\sigma) \iff a \notin \sigma \qquad (1)$$

A false pattern trace is detected when activity ordering deviates from the model:

$$FPT(\sigma) \Leftrightarrow order(\sigma) \neq order(M) \qquad (2)$$

A false decision trace occurs when the final decision differs from the standard outcome:

$$FDT(\sigma) \Leftrightarrow decision(\sigma) \neq decision_{std} \qquad (3)$$

A false throughput trace is identified when execution time exceeds or falls below expected duration:

$$FTT(\sigma) \iff |\, t_c - t_s \,| \notin \Delta_{std} \qquad (4)$$

The resulting clusters are compiled into separate event logs and passed to the analysis component, as illustrated in Figure 3.4.
The clustering component has been revised to provide a clearer formalization. Specifically, the input trace is defined as an ordered sequence of events extracted from the event log, while the expected process model is represented by the discovered Petri net. Deviation criteria are determined through conformance checking between the trace and the reference model. Four deviation categories are formally defined: Skipped Trace (ST) for missing activities, False Pattern Trace (FPT) for incorrect activity order, False Decision Trace (FDT) for incorrect decision outcomes, and False Throughput Trace (FTT) for abnormal execution times. Timing anomalies are detected using predefined temporal thresholds derived from the normal process duration. Finally, traces are clustered according to the type of deviation they exhibit, producing separate event log groups that are further analyzed in the forensic investigation phase.

3.2 Analysis Component

The analysis part looks at the identified deviations and decides which instances should be looked at first. There are three parts to it: initial analysis, time frame analysis, and track analysis (Figure 3). These subcomponents work together to provide you both numbers and pictures of questionable activity.

Figure 3 shows the analysis part of the architecture. It shows how clustered tracks are looked at more closely via initial, timing, and tracking analyses. This chart shows how the severity and rarity of deviations are used to rank high-risk instances for further in-depth forensic analysis.

The analysis component has been further clarified to strengthen the analytical explanation. The 30-day cutoff was selected based on the typical processing window observed in the dataset and domain knowledge of banking operations, where most offer-related activities are normally completed within one month. Therefore, events exceeding this period are treated as potential temporal anomalies.

The prioritization score is computed by combining two factors: the rarity of the trace variant and the severity of its deviation from the reference process model. Rarity is determined by the frequency of occurrence of each trace variant in the event log, while severity reflects the magnitude of deviation such as missing activities, incorrect ordering, or abnormal execution time. These factors are combined into a composite ranking score that allows investigators to prioritize traces with both high deviation severity and low occurrence frequency for further forensic investigation.

Figure. 3: Analysis Component diagram of proposed approach.

3.3 System Advantages

When looking at database forensics, the suggested approach could be very useful. Formal Petri net modelling clarifies processes, compliance verification and clustering aids in fraud detection, and timing and trace assessment provide in-depth behavioral analysis. Scalability and the elimination of human labor via automation allow the platform to manage enormous amounts of financial data. Changing the metrics utilized to identify inconsistencies allows the procedure to be modified to fit different business scenarios.

The purpose of this article is to provide a structured forensic framework for improving database forensics. The framework incorporates process discovery, compliance assessment, clustering, and advanced variants analysis. To modify event logs, organize deviations systematically, and rank anomalies according to rareness, timing, and traceability, we use Petri net algorithms. The following is an account of an implementation of this method that made use of process mining tools and actual financial data.

4: Implementation

This article illustrates how to utilize and evaluate the recommended model using real-world information. typical process mining tools. The documentation goes into great detail on every stage of the execution pipeline, from gathering events to performing preliminary processing, checking the legality of the process, and testing for compliance. The experiments show that the model works. We kept all of the original's figures and tables so that the scientific substance stays the same. We have shortened the tale so that readers may more easily understand the main themes of the research.

Using the ProM The tests were done using the following version of the software and main plugins: Inductive Miner, trace matching tools, and others for finding processes and checking for compliance. Before being processed, the event logs were cleaned, had their timestamps standardized, and had any incomplete records removed.

Also, the tests were also written down with the kind of CPU, quantity of RAM, and operating system used so that they may be done again. The description now includes the most important parts of the event log, such as the case ID, job title, time, and source, as well as the general characteristics of the dataset. It is also evident what rules are used to discover deviations and group them together, making the implementation description more complete and easier to understand.

5. Inductive Miner Algorithm

Inductive Miners employ recursive decomposition to construct process models that are strong and have a block structure. It could be helpful for big and complicated event logs since it assures that they will be finished correctly and won't get stuck. Figure 6 shows that the model fits to the data quite effectively and possesses very few violations

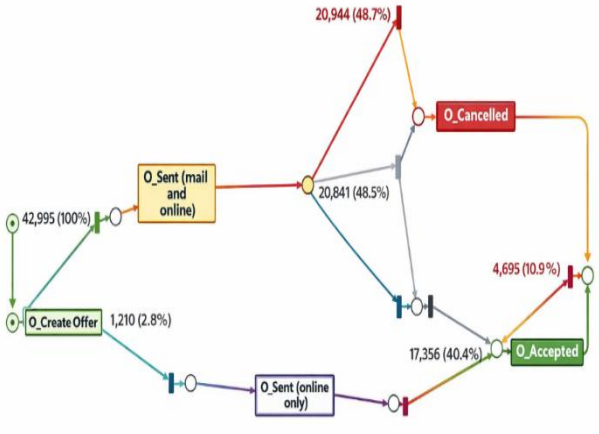

Figure 6: Discovered Process Model Using Inductive Miner Algorithm with recorded data and minimal violations.

6. Experimental Results

This page goes into great depth about the experiments that were done to see how reliable and useful the proposed model is. We tested the model in a number of situations using appropriate data and set criteria. The findings demonstrate that the model is accurate, strong, and can be used in many different circumstances. The discussion also shows that it can solve recognized issues and work better than other ways.

 ISSN: 2617-5517 | **DOI**: https://doi.org/10.59746/jfes.vxx.xx

The paper process mining method for database forensics is evaluated using common performance measures for accuracy, reliability, and computational economy. According to many sources, the accuracy measure assesses how well the identified model can avoid adding actions that aren't present in the event log. Recall is another name for the fitness statistic, which evaluates how well the model can mimic real-world behavior and spot outliers. An estimate of general accuracy is given by the F1-score, which considers both fitness and precision. In contrast to soundness, which verifies structural correctness by locating livelocks or deadlocks, execution time indicates the computing cost of constructing the model. In order to retrieve the offer lifecycle for forensic purposes, four process discovery algorithms—Alpha Miner, Heuristics Miner, inductively driven Miner, and ILP Miner—were compared using the BPI Challenge dataset. We then ran each approach with the identical parameters after preprocessing and compared the resulting models for completeness, behavioral correctness, and anomaly/suspicious pattern detection capabilities. Table 1. shows that when it comes to databases utilized in forensic investigations, Inductive Miner is the way to go because of how accurate and comprehensive the findings need to be. In addition to creating a strong, error-free model, it attained the maximum levels of precision (100%), fitness (99.60%), and F1-score (99.80%).

Table 1: Performance Comparison of Process Discovery Algorithms on the BPI Challenge Dataset

Algorithms	Fitness or Recall	Precision	F1 Score	Execution Time in Sec	Soundness
Alpha	89.90%	43.80%	58.90%	N/A	No
ILP Miner	99.9%	67.4%	80.49%	N/A	Yes
Heuristics Miner	98.10%	91.90%	94.89%	N/A	Yes
Inductive Miner	99.60%	100%	99.80%	N/A	Yes

A high F1-score of 94.89% was achieved by Heuristics Miner, which achieved a high fitness of 98.10% and precision of 91.90%. Although a little less accuracy could bring little irrelevant behavior, its sound model strikes a good compromise between the two. However, for incomplete or loud logs, it is still a dependable alternative.

Forensic inquiry is highly important, and the inductive Data Miner works far better than other technologies. A correct and dependable process model reduces the risk of incorrectly identifying deviations while making it simpler to uncover abnormalities. Forensic investigations need accurate reconstructed of event sequences; hence, structured models produced by the Inductive Miner provide a more dependable basis for detecting suspicious activities.

Alpha Miner and ILP Miner both have poor performance when compared. The incorporation of unrelated behavior is shown by ILP Miner's poor accuracy (67.40%) and high fitness (99.9%). Alpha Miner isn't up to snuff when it comes to complicated forensic situations, because to its flawed methodology and poor accuracy (43.8%). All things considered, Inductive Miner is the method that consistently produces the best results.

Results are considered "random" since tests were conducted in a non-controlled setting with fluctuating system loads and variations in optimization at the tool level. Results could not be reproduced under controlled settings; hence execution time was not included in the comparison.

6.1 Trace Analysis

This experiment analyzes activity execution sequences using trace alignment to discriminate between non-conforming and conforming process behavior. Extracted and aligned event traces allowed for the visual and analytical detection of deviations. Refer to figure 7.

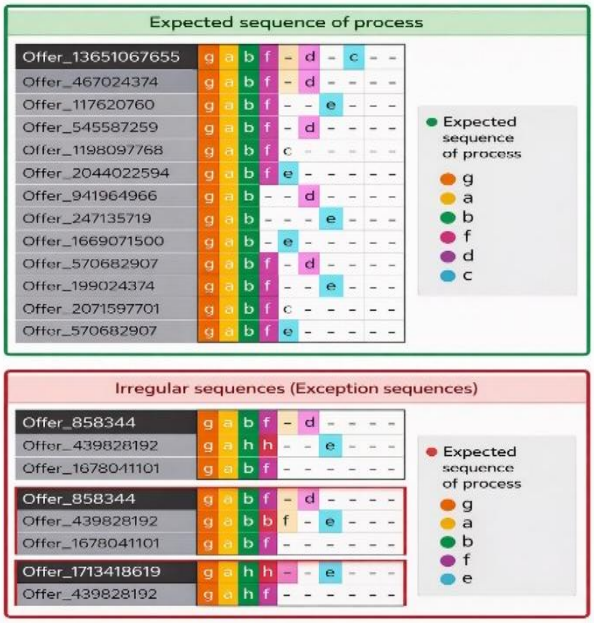

Figure 7: Trace Alignment Visualization Highlighting Conforming and Exceptional Process Variants in Offer Lifecycle Execution.

Figure 8 displays the results of the offer lifecycle trace alignment. There is one action for every row, and every row represents an instance of an offer. Rows that are consistently aligned show that the execution is running well, while placeholders indicate that activities are missing or have been rearranged.

This process is constant and regular, as shown at the top by the projected sequence of activities. There are very few deviations from the reference model. The bottom half indicates marks that don't match a reference model. Some of these variations include missing activities, phases that are out of order, and incomplete execution pathways. These reveal that the expected process behavior is not the same as the actual behavior. Some examples include sudden ends, missing steps, or events that don't follow a clear sequence. These types of alterations could suggest that the procedure isn't operating well and requires further study. But they should only be considered as probable signals of odd behavior, not evidence of fraud. Therefore, they are quite important in forensic investigations. A total of of sixteen different variations of the technique were found. Ten of the variants are spot-on with the predicted model, with six of them being particularly noteworthy. Approximately 41,549 traces follow the reference process model, while about 1,436 traces (≈3.5%) show deviations that require further forensic inspection. About 3.5% of the time, or 1,436 times, it doesn't. Consequently, forensic investigators might use trace alignment to their advantage while trying to find faults that were previously undiscovered.

 ISSN: 2617-5517 | **DOI**: https://doi.org/10.59746/jfes.vxx.xx

6.2 Timing Analysis

In order to identify commonalities and differences, this tracks the processing of loan offers across time. Every offer is represented by a row in the dotted chart, and events are shown as points over time. Outliers are highlighted after one month, but normal activity clusters appear within that time frame. The number of anomalies, the average processing time for an event, and a range of life spans are some measures. We expected the majority of offer activities to occur in the first 30 days, as shown in Figure 8. The exception to this rule is operations that extend beyond 30 days.

Temporal anomalies are activities that go beyond the anticipated processing window. Rather than providing clear evidence of fraudulent conduct, these delays are seen as forensic signs that need more examination into any aberrant operating behavior. that are brought about by humans, issues with customers, or lengthier evaluations.

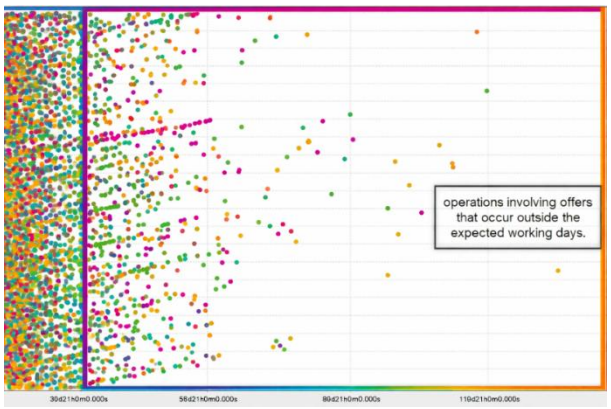

Figure 8: Dotted Chart Showing Temporal Distribution of Offer-Related Operations with Outlier Detection Beyond Expected Processing Window

Events occurring more than 30 days ago are excluded from Figure 9. In the longer timeline, many cancellations occur when refusals occur early on, while just a handful of acceptances occur after one hundred days. Unusual processing delays, such as those caused by special handling processes, operational inefficiencies, or possible policy breaches, are all indicators of late acceptance occurrences. inefficiency or even breaching the rules might be the result.

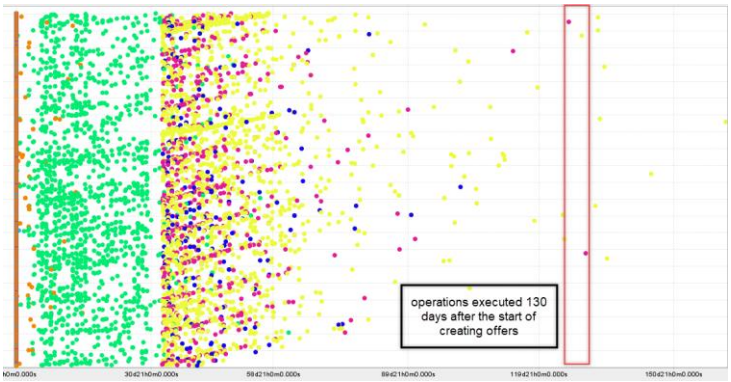

Figure 9: Temporal Distribution of Offer Outcomes beyond One Month: Highlighting Late Acceptances and Early Cancellations/Refusals

In Figure 10, we look at "O_ Accepted" events for each user. Most acceptance actions happen on time; however certain users always do late acceptances that are longer than the usual time window. This trend might indicate that something isn't being done well, that there are inefficiencies, or that something is being handled in an unusual way that needs further investigation.

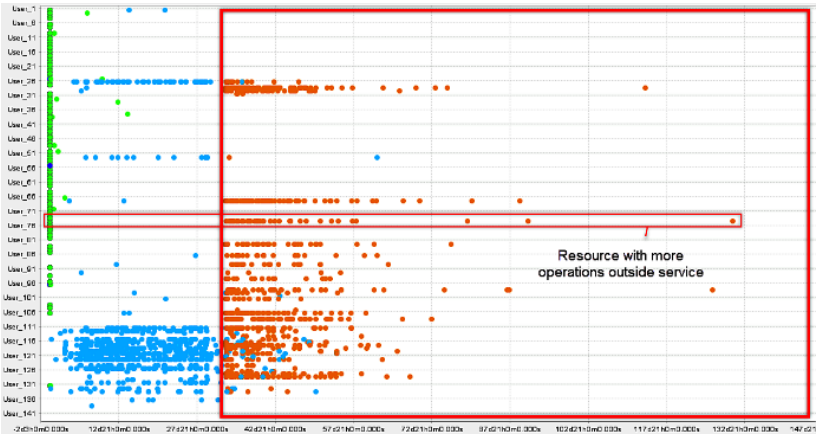

Figure 10: Temporal Distribution of 'O_Accepted' Activities per User

When a buyer waits 30 days before accepting, the transaction is considered. A total of 148 cases were identified, suggesting that delays are very substantial. Figure 11's results are supported by the fact that certain authors consistently behave in a way that is not consistent with their character.

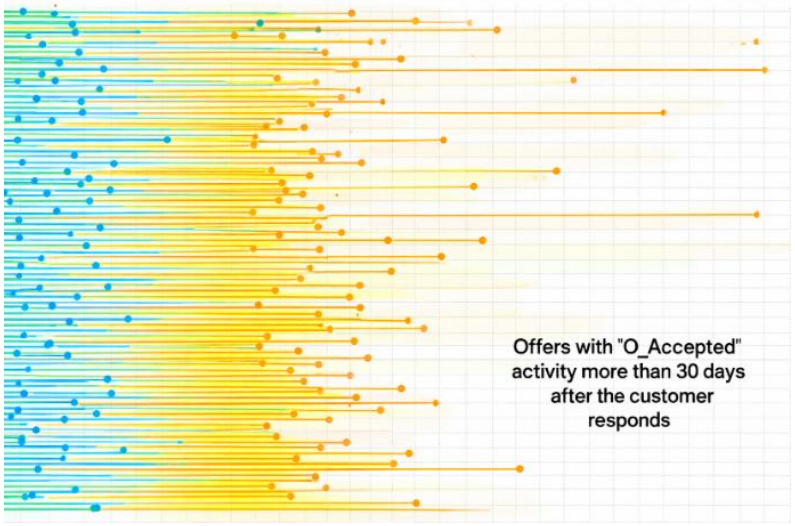

Figure 11: Analysis of Offers with "O_ Accepted" Activity Occurring Over 30 Days after Customer Response

According to timing analysis, users are doing critical actions many times, even though they were expected to execute them at different times. Communication gaps lead to delays when viewers record client replies but do not recognize them. These findings taught us a lot about operational inefficiencies and compliance issues.

since a whole, forensic investigators may benefit from both trace and timing analysis, since the former can spot structural abnormalities in activity sequences and the latter can spot unusual temporal patterns during process mining.

6.3 Comparative Analysis of Methods for Detecting Suspicious Behavior

We evaluated the suggested method's efficacy by inserting fictitious anomalies into the event logs that mimic the suspicious behaviors seen in financial systems. Missing steps, out-of-order actions, ill-considered decisions with unintended consequences, and lengthy execution times were among these issues. These variations were achieved by altering a controlled selection of songs while maintaining the integrity of the remaining information. In order to conduct controlled research on detection accuracy, around 5% of the traces was intentionally altered to reveal anomalous behavior. In this experiment, the suggested model is tested in comparison to three well-known forensic methods: ProM-based tracking [35], item set mining [36] for finding common patterns, and file-carving reconstruction [37] for detecting digital traces. To find out how effective they are, we put all of these methods to the test. The methods were tested on both actual datasets and fake logs that included additional irregularities. File-carving isn't good in structured environments, itemset mining finds frequent odd patterns without caring about a method structure, and ProM prioritizes baseline conformity above anything else.

However, the proposed approach builds a composite ranking score to highlight significant deviations, combining alignment-based conformance checking with rarity-based prioritization via the calculation of trace-variant frequencies. To evaluate the efficacy of ranking, ground-truth anomalies were paired with Spearman's correlation. The following metrics were employed for evaluation: scalability, accuracy, recall, F1-score, execution time, and prioritizing effectiveness. To ensure fair and uniform performance testing, we also conducted experiments using standardized computer equipment (Intel i7-11800H, 64GB RAM, Microsoft Windows 10). Table 2 is referenced.

Table 2: Comparative Evaluation of Forensic Process Mining Method

Method	Precision	Recall	F1-Score	Execution Time (s)	Scal-ability	Prioritization Effectiveness
ProM-Based [36]	0.71	0.78	0.74	42.1	Medium	Not Supported
Itemset Mining [37]	0.68	0.64	0.66	25.4	High	Limited
File-Carving [38]	0.59	0.53	0.56	110.7	Low	Not Applicable
Proposed Model	0.87	0.84	0.85	35.8	High	Supported (ρ = 0.82)

The proposed model outperforms ProM-based (0.74), itemset mining (0.66), and file-carving methods (0.56), as shown in Table 2, achieving the best F1-score of 0.85. With alignment-based deviation detection and rarity-based ranking combined, the number of true positives and the number of false positives may be significantly reduced. In addition, the suggested model is efficient; it handles huge logs in 35.8 seconds, which is far quicker than carving files (110.7 seconds) and on par with itemset mining, all while providing better accuracy. The prioritization

mechanism was evaluated using Spearman rank correlation, which measures the agreement between the computed prioritization score and the ground-truth anomaly severity. The obtained correlation value of 0.82 indicates a strong alignment between the proposed ranking method and the actual importance of suspicious cases, The ground truth was defined based on the traces that were intentionally modified during the anomaly injection process. These traces were labeled as anomalous, while the remaining unmodified traces were considered normal. This labeling enabled objective evaluation of anomaly detection performance.

Common classification measures were used to assess the detection performance. The F1-score is the harmonic average of the two metrics, recall and precision, which are used to determine the percentage of real anomalies discovered and the fraction of anomalies that are truly abnormal.

Scalability was assessed based on the ability of each method to process increasing event log sizes within reasonable computational time and memory usage. Methods capable of efficiently handling large logs were classified as High, moderately efficient methods as Medium, and approaches with significant computational overhead as low. it facilitates successful case prioritizing and is different from baseline techniques. For practical forensic auditing, this means it is accurate and can be used. The suggested model is robust and has forensic relevance, according to the tests. When it came to creating robust process models, Inductive Miner stood head and shoulders above the competition as the best discovery method. About 3.5% of the traces that needed more inquiry were detected by successful trace alignment. Processing bottlenecks, delays, and user-specific anomalies were uncovered by temporal analysis. Lastly, the results of the comparative assessment showed that the efficiency and accuracy of anomaly identification are much enhanced when conformance checking is combined with rarity-based prioritizing. Taken as a whole, these findings prove that the model is an all-inclusive and efficient tool for forensic process evaluation.

7. Conclusion

This paper suggested a forensic framework based on process mining for finding suspicious behaviors in databases of financial transactions. The framework combines process discovery, compliance verification, behavioral grouping, and anomaly prioritizing to make it easier to automatically investigate transactional activity. The suggested technique was shown to work well in an experiment utilizing the BPI Challenge database and the ProM mining process environment. The Inductive Miner approach was the best in making a strong process model, with an F score of 100% and an F1-score of 99.8%. We also found behavioral and temporal abnormalities by looking at the position and timing of the traces. This showed that about 3.5 percent of the tracks exhibited patterns that seemed suspicious and warranted additional forensic investigation. The proposed approach had the highest F1-score (0.85) and did a good job of putting questionable cases first. This suggests that it works better than baseline techniques for finding anomalies. These results are promising, however there are several problems with the recommended strategy. The first is that in real-world banking networks, the quality and completeness of event logs might be different, which would have a big effect on how well the analysis works. Second, the trials only employed a tiny amount of data, therefore the conclusions may not be true in different cases. Additionally, sensitivity in diverse operational contexts may be influenced by designated threshold values used for time-based anomaly detection. Possible future directions for research include testing the system on more extensive and varied financial datasets, making it more robust and scalable in real-world forensic settings by incorporating machine learning methods for adaptable anomaly detection, and creating automated methods for selecting thresholds.

Acknowledgments

The authors extend their sincere thanks and appreciation to everyone who contributed to the completion of this research, and they also express their gratitude to the University of Wasit / College of Computer Science and Information Technology for providing a supportive academic environment.

Conflict of Interest

The authors declare that they have no conflicts of interest.

Data Availability

The data used in this study are publicly available through the BPI Challenge 2017 dataset, and their provenance has been documented and referenced within the body of the paper, ensuring that the results can be verified and reproduced.

References

[1] Srivastava, S., & Bhatnagar, R. (2021). Process mining techniques for detecting fraud in banks: A study. *Turkish Journal of Computer and Mathematics Education*, *12*(12), 3358-3375.

[2] Olivier, M. S. (2009). On metadata context in database forensics. *Digital Investigation*, *5*(3-4), 115-123.

[3] Cankaya, E. C., & Kupka, B. (2016, December). A survey of digital forensics tools for database extraction. In *2016 future technologies conference (ftc)* (pp. 1014-1019). IEEE.

[4] Van Der Aalst, W. M., Reijers, H. A., Weijters, A. J., van Dongen, B. F., De Medeiros, A. A., Song, M., & Verbeek, H. M. (2007). Business process mining: An industrial application. *Information systems*, *32*(5), 713-732.

[5] Erdogan, T. G., & Tarhan, A. (2018). Systematic mapping of process mining studies in healthcare. *IEEE Access*, *6*, 24543-24567.

[**6**] Sundari, M. S., & Nayak, R. K. (2020). Process mining in healthcare systems: a critical review and its future. *International Journal of Emerging Trends in Engineering Research*, *8*(9).

[**7**] Cerezo, R., Bogarín, A., Esteban, M., & Romero, C. (2020). Process mining for self-regulated learning assessment in e-learning. *Journal of Computing in Higher Education*, *32*(1), 74-88.

[**8**] Céu, H., Grilo, C., Rijo, R., & Martinho, R. (2024). Mining resource usage in molds manufacturing processes through process mining. *Procedia Computer Science*, *239*, 2359-2368.

[**9**] Castiglione, C. (2024). Automated generation of digital models for manufacturing systems: The event-centric process mining approach. *Computers & Industrial Engineering*, *197*, 110596.

[1**0**] Maita, A. R. C., Martins, L. C., López Paz, C. R., Rafferty, L., Hung, P. C., Peres, S. M., & Fantinato, M. (2018). A systematic mapping study of process mining. *Enterprise Information Systems*, *12*(5), 505-549.

[1**1**] van der Aalst, W. M., Bolt, A., & van Zelst, S. J. (2017). RapidProM: mine your processes and not just your data. *arXiv preprint arXiv:1703.03740*.

[1**2**] Adriansyah, A., van Dongen, B. F., & van der Aalst, W. M. (2011, August). Conformance checking using cost-based fitness analysis. In *2011 ieee 15th international enterprise distributed object computing conference* (pp. 55-64). IEEE.

[1**3**] Jans, M., Alles, M. G., & Vasarhelyi, M. A. (2014). A field study on the use of process mining of event logs as an analytical procedure in auditing. *The Accounting Review*, *89*(5), 1751-1773.

[1**4**] dos Santos Garcia, C., Meincheim, A., Junior, E. R. F., Dallagassa, M. R., Sato, D. M. V., Carvalho, D. R., ... & Scalabrin, E. E. (2019). Process mining techniques and applications–A systematic mapping study. *Expert Systems with Applications*, *133*, 260-295.

[**15**] Macak, M., Daubner, L., Sani, M. F., & Buhnova, B. (2022). Process mining usage in cybersecurity and software reliability analysis: A systematic literature review. *Array*, *13*, 100120.

[**16**] Rojas, E., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. (2016). Process mining in healthcare: A literature review. *Journal of biomedical informatics*, *61*, 224-236.

[**17**] Atlam, H. F., Ekuri, N., Azad, M. A., & Lallie, H. S. (2024). Blockchain forensics: A systematic literature review of techniques, applications, challenges, and future directions. *Electronics*, *13*(17), 3568.

[**18**] Igonor, O. S., Amin, M. B., & Garg, S. (2025). The application of blockchain technology in the field of digital forensics: A literature review. *Blockchains*, *3*(1), 5.

[**19**] Khalid, Z., Iqbal, F., & Saqib, M. (2025). Bridging knowledge gaps in digital forensics using unsupervised explainable AI. *Forensic Science International: Digital Investigation*, *53*, 301924.

[**20**] Felix, A. O. (2025). Enhancing Digital Forensics Investigations Using AI Driven Anomaly Detection and Log Correlation: A Mixed Methods Approach. *International Journal of Future Engineering Innovations*.

[**21**] Vyas, D., Shah, M., Kothari, A., Golakia, J., & Parikh, V. (2025). Enhancing Digital Forensics: Machine Learning Techniques for Social Media Investigation. *Procedia Computer Science*, *258*, 2290-2301.

[**22**] Andersen, D. B., Sunde, N., & Porter, K. (2025). Tool induced biases? Misleading data presentation as a biasing source in digital forensic analysis. *Forensic Science International: Digital Investigation*, *52*, 301881.

[**23**] Hargreaves, C., van Beek, H., & Casey, E. (2025). SOLVE-IT: A proposed digital forensic knowledge base inspired by MITRE ATT&CK. *Forensic Science International: Digital Investigation*, *52*, 301864.

[**24**] Kim, K. J., Lee, C. H., Bae, S. E., Choi, J. H., & Kang, W. (2025). Digital forensics in law enforcement: A case study of LLM-driven evidence analysis. *Forensic Science International: Digital Investigation*, *54*, 301939.

[**25**] Selim, A., & Ali, İ. (2024). The role of digital forensic analysis in modern investigations. *Journal of Emerging Computer Technologies*, *4*(1), 1-5.

[**26**] Mpungu, C., George, C., & Mapp, G. (2024). Digital Forensics Readiness in Big Data Networks: A Novel Framework and Incident Response Script for Linux–Hadoop Environments. *Applied System Innovation*, *7*(5), 90.

[**27**] Choi, H., & Lee, S. (2023). Forensic analysis of SQL server transaction log in unallocated area of file system. *Forensic Science International: Digital Investigation*, *46*, 301605.

[**28**] Mohamed, R. A., & Kassem, G. (2023, October). Development of Conceptual Model for Performing Process Mining on Blockchain Data: A Cybersecurity Approach. In *2023 2nd International Conference on Smart Cities 4.0* (pp. 174-178). IEEE.

[**29**] Van der Aalst, W. M. (2022). Process mining: a 360 degree overview. In *Process mining handbook* (pp. 3-34). Cham: Springer International Publishing.

[**30**] Khan, A. A., Shaikh, A. A., Laghari, A. A., Dootio, M. A., Rind, M. M., & Awan, S. A. (2022). Digital forensics and cyber forensics investigation: security challenges, limitations, open issues, and future direction. *International Journal of Electronic Security and Digital Forensics*, *14*(2), 124-150.

[**31**] Aljahdali, A. O., Alluhaib, G., Alqarni, R., Alsharef, M., & Alsaqqaf, A. (2022). Big data analysis and forensics. *International Journal of Electronic Security and Digital Forensics*, *14*(6), 579-593.

[**32**] Tariq, Z., Charles, D., McClean, S., McChesney, I., & Taylor, P. (2022). Anomaly detection for service-oriented business processes using conformance analysis. *Algorithms*, *15*(8), 257.

[**3**3] van Zelst, S. J., Mannhardt, F., de Leoni, M., & Koschmider, A. (2021). Event abstraction in process mining: literature review and taxonomy. *Granular Computing*, *6*(3), 719-736.

[**3**4] Broer Bahaweres, R., Trawally, J., Hermadi, I., & Imam Suroso, A. (2021, February). Forensic audit using process mining to detect fraud. In *Journal of Physics: Conference Series* (Vol. 1779, No. 1, p. 012013). IOP Publishing.

[**3**5] Rodríguez-Quintero, J. F., Sánchez-Díaz, A., Iriarte-Navarro, L., Maté, A., Marco-Such, M., & Trujillo, J. (2021). Fraud audit based on visual analysis: A process mining approach. *Applied Sciences*, *11*(11), 4751.

computer networks (ISCON) (pp. 584-590). IEEE.

[36] R. Broer Bahaweres, J. Trawally, I. Hermadi, and A. Imam Suroso, "Forensic Audit Using Process Mining to Detect Fraud," Journal of Physics: Conference Series, vol. 1779, no. 1, 012013, pp.1-11, 2021.

[37] S. Khan, S. Parkinson, and C. Murphy, "Context-based irregular activity detection in event logs for forensic investigations: An itemset mining approach," Expert Systems with Applications, vol. 233, 120991, pp.1-13, 2023.

[38] L. Englbrecht, S. Schönig, and G. Pernul, "Supporting Process Mining with Recovered Residual Data," In the Practice of Enterprise Modeling, pp. 389-404. Cham, Springer International Publishing, 2020.