RUTGERS
Rutgers Business School
Newark and New Brunswick

universiteit hasselt

PROCESS MINING OF EVENT LOGS IN AUDITING: OPPORTUNITIES AND CHALLENGES

Mieke Jans

Hasselt University

Michael Alles
Miklos Vasarhelyi

Rutgers Business School

What is Process Mining of Event Logs?

- *The basic idea of process mining is to extract knowledge from **event logs** recorded by an information system. Until recently, the information in these event logs was rarely used to analyze the underlying processes. Process mining aims at improving this by providing techniques and tools for discovering process, control, data, organizational, and social structures from event logs. Fuelled by the omnipresence of event logs in transactional information systems… process mining has become a vivid research area.*
- http://is.tm.tue.nl/staff/wvdaalst/BPMcenter/process%20mining.htm

Process Mining and Auditing

- Event Logs also referred to as Audit Trails.
- **Audit trail**: *a chronological record of computer system activities which are saved to a file on the system. The file can later be reviewed by the system administrator to identify users' actions on the system or processes which occurred on the system.*
- Despite the presence of the word "audit" in the term audit trail, it does not refer to auditing in the accounting sense, but to the general potential an event log provides to reconstruct past transactions.
- Little use made by auditors of process mining to examine the data contained in event logs.

Recall World of Manual Accounting

- In a world of manual bookkeeping the data which auditors had to rely on when checking what transactions the client firm had undertaken and how it had accounted for those transactions came entirely from paper based ledgers.
- If ledger entries have been falsified, erased and overwritten, added to or modified at another date and time by the same or other party, auditors can detect that only through physical scrutiny of the books.
- Similar to bankers having to rely on a bank teller's familiarity with the signature of a customer to determine whether a check was genuine or forged.

The WYSIWYG Problem in Auditing

- Hand written ledgers suffer from what we call the "what you see is what you get" or WYSIWYG problem: the only information that the auditor has is what they can literally observe in front of them.
- Hence, the auditor has no way of verifying who made those ledger entries and when they did so.
- What makes an event log such a unique and potentially invaluable resource for auditing is not only that it provides the auditor with more data to analyze, but also because that additional data is recorded automatically and independently of the person whose behavior is the subject of the audit.

Meta-Data in Event Logs

- With an event log the auditor is no longer restricted to the WYSIWYG ledger of transactions entered by the auditee, but also possesses an independent set of "meta-data" about the circumstances under which the auditee made those entries.
- That meta-data encompasses more than simple time stamps for transactions, for by taking advantage of that tracking data, the event log enables the auditor to reconstruct the history of any given transaction.
- Hence, the auditor is now able to trace the relationship of that particular entry and its author to all prior transactions by that or related parties.

The Digital Economy Facilitates Process Mining and Event Log Creation

- The emergence of the digital economy has fundamentally altered both the way of running businesses and of performing [continuous] audits.
- Most businesses of any significant size today store their data electronically thanks to the maturing of technologies for databases and computer networks.
- Systems such as Enterprise Resource Planning (ERP), Workflow Management (WFM), Customer Relationship Management (CRM), Supply Chain Management (SCM) or Busines-to-Business (B2B) all create and store histories that potentially can be process mined by auditors.

Process Mining Still More Potential than Reality

- It is important to understand that the term event "log" promises more than what is the reality today: the event log is a database of time, process, and originator stamps, not a chronicle that an auditor can simply read as if it were a story.
- While most modern IT systems—especially ERP systems—record a history, that is the starting point of creating an event log, not one in itself.
- Extracting an event log from an ERP database requires considerable manual effort by the auditor.
- Logging can slow down ERP systems, so often that feature is turned off, especially if no one wants it.

An Example of An Event Log of an Invoice

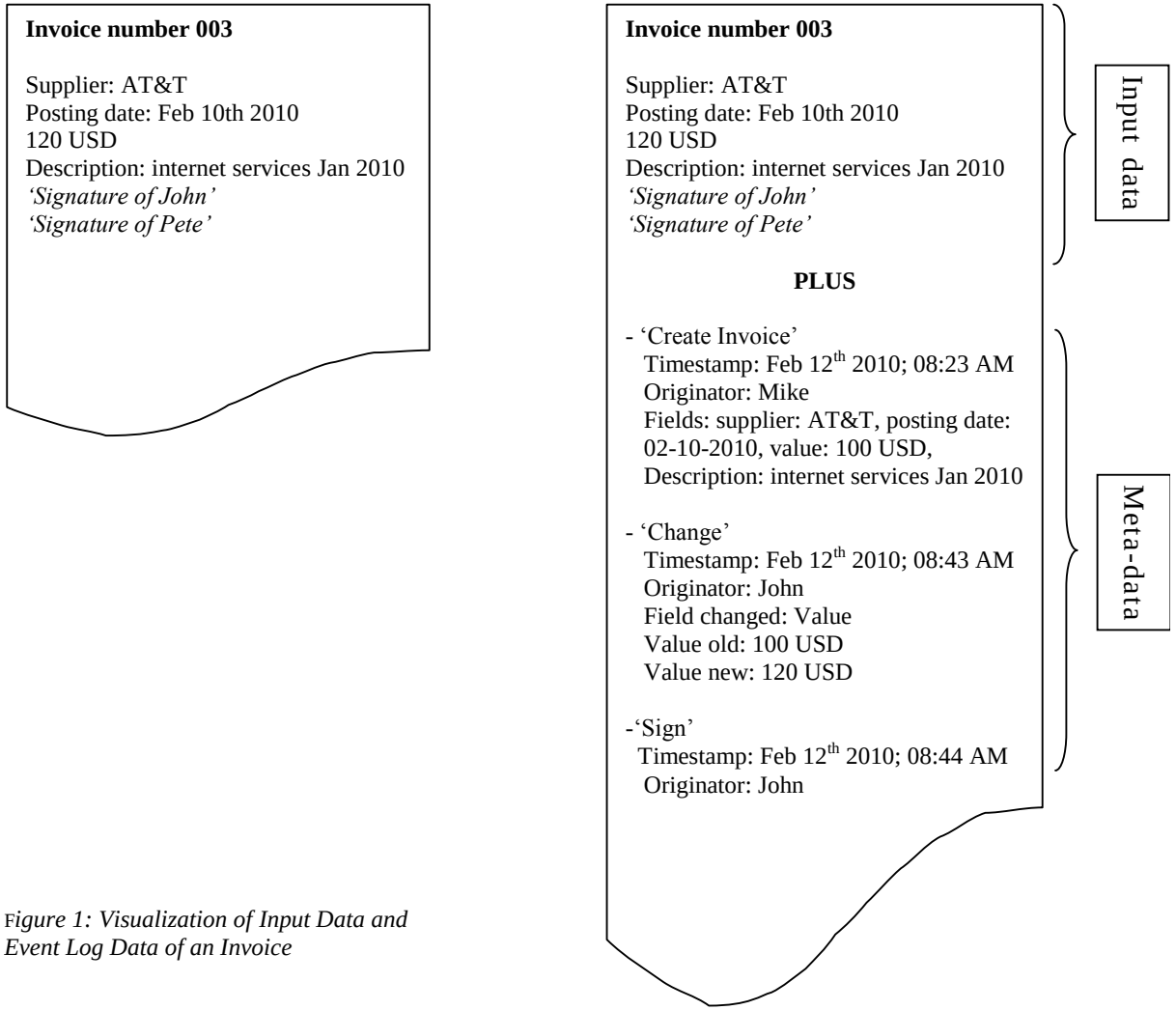

Figure 1: Visualization of Input Data and Event Log Data of an Invoice

Components of an Event Log: Input Data

- The left hand side of Figure 1 shows the data about the invoice that is entered by the person making the data entry into the firm's ledger.
- This is what we call **input data** as it is characterized by the controllable act of a person inputting the data.
- This input data is the type of data available at the moment this data is stored, such as the invoice number, the posting date, the supplier etc.
- This is also the type of transactional data that is currently used for auditing and monitoring.

Components of an Event Log: Meta Data

- The right hand side of Figure 1 shows the data that is stored in the event log of that same invoice.
- Formally speaking all input data is also part of the event log, but clearly it is the entries that are unique to the event log that are of particular interest to an auditor because that data is recorded automatically by the system and not inputted by the auditee.
- It is this meta-data which makes an event log of larger dimension
- That contextual meta-data enables the auditor to identifying relationships between this transaction and all other transactions in the database.

Comparing Input Data with Meta Data

- Event log reveals changes to the invoice and the identity of other individuals who "touched" the invoice in any way during its progress through the purchase to pay business process.
- For example, where we see a posting date of February 10th in the input data, this could differ from the system's timestamp recorded in the event log.
- The meta-data in the event log can potentially contain many pieces of information, like at what times which fields are changed by which originators.
- Meta data serves as a check on the input data.

Using Process Analysis of Meta Data to Create a Narrative History of the Business Process

- Meta-data, when combined with the input data, enables the auditor to reconstruct the history of a particular transaction.
- Thus, the following activities can be reconstructed from the data shown in the event log example of Figure 1:

1. on Feb 12, 8:23 AM: Mike entered invoice No. 3 in system, filling out the supplier (AT&T), posting date (02-10-2010), invoice value (100 USD) and description (internet services Jan 2010)
2. on Feb 12, 8:43 AM: John changed 'Value' from '100USD' to '120USD'
3. on Feb 12, 8:44 AM: John signed invoice No. 3

Meta Data Expands Audit Evidence

- Where auditors used to have only the paper-written general ledgers with the data on it that the auditee (or anyone else) had written down, there is now event log data available that records everything that happened to that ledger.
- For example, where the auditor used to rely on the signature at the bottom of a document to verify who was responsible for the document, the event log keeps track of, amongst other things, the person opening/altering/signing the document.
- Critically for auditing, meta data in event log is independent of the person entering the input data.

Process Mining Overcomes WYSIWYG Problem

- Thus the event log enable auditors to overcome the WYSIWYG constraint of the input data.
- This event log forms the start and also the opportunity to enact process mining.
- Based on the log of these events, a process model can be used to gain insights about what *actually happens* in a process, in contrast to what people *think* happens in a process.
- Many other such comparisons can be undertaken by the systematic analysis of a properly constructed, comprehensive event log.
- Key is creating such a log in the first place.

Steps in Creating an Event Log 1

- The event log data that is captured by the ERP system is large and dispersed over numerous tables.
- In order to mine the event log and, hence, the process, a rigorous and defensible method of structuring the data needs to be developed.
- The first step is for the auditor to develop a holistic understanding of the activities that constitute the process being audited, trading off comprehensiveness against the size of the resulting event log.
- The second step in event log creation is the selection of the *process instance* which undergoes the identified activities, such as an invoice.

Steps in Creating an Event Log 2

- The activity performed by the auditee is input data but the storage of the act itself is meta-data because it is beyond the control of the auditee.
- Each process instance has a unique identifier and is stored along with attributes that describe the process instance, like for instance the size, the supplier involved etc.
- The activities performed on the process instance, called *unique event entries*, also have a unique identifier which is stored along with the timestamp and originator stamp.

Steps in Creating an Event Log 3

- Aside from the time and originator stamps, other attributes that describe the event or activity are stored. All these attributes are mainly stemming from meta-data.
- Which activities and attributes will be captured in the event log is jointly determined by capabilities of the ERP system to record data and the judgment of the auditor as to the necessary scope of the event log.

Example of a Constructed Event Log

Unique Event Entries (UEE)				
UEE-ID	**PI-ID**	**Activity**	**Originator**	**Timestamp**
UEE -ID 1	PI-ID 1	Activity A: Change PO	Originator X	xx/xx/xxxx
UEE -ID 2	PI-ID 1	Activity B: Enter Goods Receipt	Originator Y	xx/xx/xxxx
UEE -ID 3	PI-ID 1	Activity A: Change PO	Originator Z	
UEE -ID 4	PI-ID 2	…	…	
…				
UEE -ID m	PI-ID n			

UEE attributes		
UEE -ID	**Name**	**Value**
UEE -ID 1	field changed	delivery address
UEE -ID 1	old value	'previous delivery address'
UEE ID 2	Goods Receipt number	GR0005014
UEE -ID 2	reference to invoice	SP14V51
UEE -ID 3	field changed	commercial discount %
…		
UEE -ID m	…	…

Process Mining as an Audit Tool

- In auditing the information that is analyzed to issue an audit opinion is essentially the same as when an audit was performed in a paper-based setting: input data, albeit now in a database and not a paper ledger.
- This data may now be analyzed using search queries rather by manual examination, but that is simply automating an existing manual procedure.
- Process mining is reengineering audit practice to take full advantage of the capabilities of digital businesses.
- Using event logs the auditor can examine not just what has happened, but also who did it, when, and with whom, as well as similar transaction elsewhere.

Process Mining Based Audit tests 1

- Fraud typically takes place at times where there are fewer other employees around to ask questions, such as at lunchtime.
- Hence, forensic accountants monitor the firm's ledgers at lunchtime to see who else is on the system, whereas the person committing the fraud might wish to cover their tracks by entering a different time for when they undertook the fraudulent transaction.
- An event log will record the actual time of this transaction and that information is available to the auditor without recourse to monitoring at precisely the same time as the data is being entered.

Process Mining Based Audit tests 2

- An anomaly that could be revealed by mining the event log data and which may not be detected so readily through other means is violations of segregation of duty controls.
- Some SOD violations are detectable with standard audit tests.
- But employees not following required procedures like first getting approval and only then ordering the goods would not be apparent by only looking at transactional data alone.
- Using the timestamps in the event log makes visible this circumvention of procedures.

Process Mining Based Audit tests 3

- Consider a collaboration between a supplier and an employee, systematically changing a purchase order within the sustained margins after a last approval to this purchase. The supplier gets paid more than was agreed upon, and can provide kickbacks to the involved employee.
- This abuse can be discovered by analyzing event log data since it captures changes to the invoice. Because of the stored originators of activities it is also possible to analyze collaboration between employees and other parties.
- Compensates for weaknesses in ERP controls.

Process Mining as an Analogy to Security Cameras (CCTV) 1

- An analogy that can be applied to event logs is that of video surveillance used in businesses to safeguard assets and deter crime.
- The major difference between event logs and video surveillance recordings is that storing transactional data is cheap and that it is time and location stamped so that it is feasible for an auditor to search for anomalies and track back history in the event of a detected problem, such as theft or fraud.
- By contrast, many surveillance cameras are actually non-functioning replicas since there is no possibility of cost effectively monitoring their feeds.

Process Mining as an Analogy to Security Cameras (CCTV) 2

- Firms install fake cameras because they have a deterrent effect, which is also the greatest benefit from known systematic process mining of event logs.
- Just the chance that someone might be watching a person's behavior can serve to constrain that behavior.
- How much more effective this deterrence effect would be then if an auditee knew that event logs were indeed being automatically and continuously monitored for anomalies and subject to tests of analytic procedures?

Methods of Process Mining in Auditing 1

- Three fundamental process mining *perspectives*: the process perspective ("How?"), the organizational perspective ("Who?") and the case perspective ("What?").
- The *process perspective* uses the time and location stamps in the event log to help answer the question of "How the process was undertaken?" with respect to the following order of activities.
- This perspective can be used by auditors to compare the process as it is meant to be performed against how it actually is and thus identify control failures and weaknesses.

Methods of Process Mining in Auditing 2

- The *organizational perspective* uses the data in the 'Originator' field in the event log to help answer the question of "Who was involved in the process?"
- In this perspective, underlying relations between performers or between performers and tasks can be made visible.
- The obvious use of this perspective in auditing is in checking segregation of duty controls, either retrospectively, to check existing procedures, or prospectively, to verify integrity of controls when personnel changes are expected (for example, due to layoffs or expansion).

Methods of Process Mining in Auditing 3

- The *case perspective* or the "What happened with this particular transaction?" question focuses on a single case, tracing back its history and relationships of parties that are involved in that history.
- This perspective can be used to check assertions on a case basis, for instance "When a purchase order is changed by over 2%, is this followed by a new approval?". (Assuming that the relative magnitude of a change is stored as a attributes in the event log).
- This will also be useful to analyze the separately stored attributes, for example the size of an order or the related supplier.

Methods of Process Mining in Auditing 4

- Another way of classifying process mining is by the approach followed to search for answers to these three questions.
- Broadly speaking, there are at least five different such *tasks* in process mining: a. process discovery, b. conformance check, c. performance analysis, d. social networks analysis, e. decision mining and verification.
- The application of only some of these tasks to auditing have been examined in depth.
- Need for much more research on how the data contained in event logs can be exploited through process mining.

Caveats and Future Research

- It is essential that the integrity of the meta-data is assured, so that it is beyond the influence of the auditee. Otherwise the value of the event log as audit evidence is lost.
- This is a "guarding the guards" situation with meta-controls necessary.
- The format of the event log is still being intensively researched, and it is important that auditors follow developments in this areas closely in order not to fall behind other fields that use process mining.
- Similarly, new developments are arising in the methodology of process mining.

Next Steps

- Many of the examples of process mining as applied to auditing that we present in this paper take as their starting point familiar manual audit procedures.
- Thus, we are following the standard route in technology adoption, which is to first automate manual processes and only then, once a level of comfort is attained, to reengineer those processes to take full advantage of the capabilities of the technology
- Once the potential of process mining as applied to auditing is recognized, its true value added can be fully investigated.