Thomas Stocker, Rafael Accorsi, Tobias Rother

Computergestützte Prozessauditierung mit Process Mining

Existierende Mechanismen zur Prozessauditierung reichen für einen wirksamen Nachweis der Einhaltung verbindlicher Vorgaben und interner Richtlinien nicht aus. Process Mining Verfahren können in diesem Zusammenhang dazu beitragen, die Verlässlichkeit und Aussagekräftigkeit von Analyseergebnissen durch Anwendung von Massendatenanalysen erheblich zu steigern. Dieser Artikel zeigt am Beispiel eines Einkaufsprozesses die Chancen und derzeitigen Grenzen dieser im Prüfungswesen stark an Bedeutung gewinnenden Technologie.

Kolumnentitel:
Prozessauditierung mit Process Mining

Stichwörter:
Geschäftsprozess, Prozessauditierung, Prozessanalyse, Process Mining, IT-Prüfung, IT-Revision, Massendatenanalyse

1 Transparenzverluste in automatisierten Geschäftsprozessen

Technologische Entwicklungen im Bereich betrieblicher Informationssysteme eröffnen Unternehmen vielfältige Möglichkeiten zur effizienten Steuerung interner Abläufe. Andererseits birgt die IT-gestützte (teil-)automatisierte Ausführung von Geschäftsprozessen auch Risiken. Vor dem Hintergrund stetiger Optimierungsbestrebungen erfordert die Steuerung von Prozessen die Berücksichtigung verbindlicher gesetzlicher, datenschutzrechtlicher und regulatorischer Vorgaben, die im Falle der Nichtkonformität kostspielige Sanktionen mit sich bringen können. Die technische Implementierung von Prozessen erfordert die Kopplung diverser Teilsysteme (CRM, ERP, BPM) und führt bei mangelnder Implementierung zu einem Transparenzverlust, der gezielte Prozessmanipulationen, aber auch unbeabsichtigte Konfigurationsfehler begünstigt und sich dabei gleichzeitig negativ auf die Auditierbarkeit von Prozessen auswirkt [Accorsi 2011]. Als Folge bleiben Betrugsfälle oder Schwachstellen [Lowis 2011] oft lange unentdeckt und richten erheblichen Schaden an. Gesetzliche Vorgaben, die der Qualitätssicherung von Unternehmensprüfungen dienen (SOX-Act, EU-Richtlinie 2006/43/EG), haben an der Situation wenig verändert, wie aktuelle Betrugsfälle bei der französischen Société Générale oder der schweizer UBS belegen. Über die Ursachen kann in vielen Fällen nur spekuliert werden. Plausibel scheint jedoch, dass mit der durch den Wunsch nach größtmöglicher Flexibilität bei der Prozessausführung motivierten Unterlassung bzw. Abmilderung notwendiger Kontrollen in Bezug auf Compliance und Sicherheit Risiken eingeführt oder begünstigt werden, denen im Rahmen von ex-post Prüfungen nur unzureichend begegnet werden kann. Die Hauptursache hierfür liegt in der Komplexität von Prozessaudits, die neben einem umfassenden Verständnis des vorgesehenen Ablaufs die Prüfung tatsächlich aufgetretenen Prozessverhaltens anhand „digitaler Spuren“ erfordert. Während Aussagen auf Basis existierender Prozessdokumentation prinzipiell wenig

aussagekräftig sind, wird bei der Verhaltensanalyse meist auf Stichprobenverfahren und simulierte Testläufe zurückgegriffen [Carlin 2007]. Tiefgreifende Analysen kommen aufgrund der kostenintensiven Natur manueller Prüfungen oft nicht in Frage. Die dadurch implizierte inhärente Unvollständigkeit liefert eine nachvollziehbare Begründung für unentdecktes Fehlverhalten von Geschäftsprozessen.
Process Mining bietet Verfahren zur Analyse von Prozessdaten. Insbesondere kann dabei nicht-konformes Prozessverhalten entdeckt und so das Augenmerk eines Prüfers auf die aus Analysesicht interessanten Stellen gelenkt werden [Accorsi 2012]. Dieser Beitrag stellt Konzepte und Werkzeuge des Process Mining vor und veranschaulicht dessen Einsatzfähigkeit in Prozessprüfungen anhand eines Fallbeispiels.

2 Process Mining

Process Mining [van der Aalst 2011] ist eine dem Data Mining entstammende wissenschaftliche Disziplin, die sich der Analyse prozessorientierter Vorgänge widmet. Traditionell befasst sich ein Großteil der Arbeiten mit der Rekonstruktion von Prozessmodellen aus Logdateien [van Dongen 2009] und dem Vergleich von existierenden Modellen mit aufgezeichnetem Prozessverhalten [Cook 1999]. Eine Anwendung auf die Prüfung von Geschäftsprozessen ist deshalb naheliegend [van der Aalst 2010]. Process Mining stellt in diesem Zusammenhang die Verbindung zwischen der umfangreichen Datenerhebung und -speicherung im Sinne von Big Data und dem Geschäftsprozessmanagement (BPM) dar. Neben detaillierten statistischen Informationen über analysierte Prozesse, die vor allem für die Prozessoptimierung relevant sind, stellt Process Mining in drei verschiedenen Bereichen Methoden zur Verfügung, die im Rahmen von Prozessaudits zur Evaluierung der tatsächlich erreichten Prozesskonformität und -performanz dienen können:

1. **Prozesserkennung** (Process Discovery)
   Auf Basis von aufgezeichnetem Prozessverhalten wird ein strukturiertes Prozessmodell extrahiert, das die Abfolge von Prozessaktivitäten in einer kompakten Form repräsentiert und Information über Verzweigungen und Schleifen enthält. Diese Form der Analyse lässt Rückschlüsse darauf zu, wie der Prozess tatsächlich „gelebt“ wird.

2. **Prozesskonformität** (Conformance Checking)
   Ausgehend von einem Prozessmodell für das idealisierte Prozessverhalten oder einer mathematischen Beschreibung einzuhaltender Anforderungen wird überprüft, welche Abweichungen das aufgezeichnete Prozessverhalten aufweist. Somit lässt sich feststellen, ob vorgeschriebene Prozessschritte (z.B. zur Durchsetzung von Freigaberichtlinien) umgangen oder Anforderungen bzgl. der Funktionstrennung (z.B. 4-Augen-Prinzip) eingehalten wurden.

3. **Prozesserweiterung** (Enhancement)
   Motiviert durch das Bestreben, eine optimierte Variante eines vorliegenden Prozesses abzuleiten, wird dessen Performance im Hinblick auf zeitliches Verhalten (Durchlaufzeiten) oder Ressourcenauslastung untersucht, um Teilprozesse zu identifizieren, die sich negativ auf die Gesamtperformanz des Prozesses auswirken.

Die Voraussetzung für die Anwendbarkeit von Process Mining ist eine prozessorientierte Sichtweise auf Daten. Die Ereignisse innerhalb Prozesslogs müssen dabei Aktivitäten im zugrundeliegenden Prozessmodell zuordenbar sein. Jedes Ereignis ist zudem einem Fall (einer

Prozessinstanz) zugeordnet der für einen konkreten Ablauf des Prozesses steht. Zusätzliche Informationen über Aktivitäten können Angaben zu ausführenden Subjekten, den Zeitstempel der Ausführung oder Charakteristika im Prozess verarbeiteter Datenobjekte umfassen (vgl. Tab. 1).

Zeitstempel	Instanz	Aktivität	Akteur	Attribut
2012-09-01 10:30:45	1	Bestellung anlegen	User 3544	Bestellnr.: 204432
2012-09-10 14:33:12	1	Wareneingang	User 1282	Belegnr.: 239431237

Tab. 1: Schematische Darstellung eines Prozesslogs

Die Sammlung und Konsolidierung prozessrelevanter Information muss in einem vorgelagerten Datenerhebungsprozess stattfinden. Dieser Prozess ist vom Unternehmenskontext und den Anforderungen an die Analyse abhängig. Während mittlerweile Informationssysteme mit expliziter Prozessorientierung (PAIS) am Markt erscheinen, die den Ablauf von Geschäftsprozessen anhand zuvor definierter Prozessmodelle steuern [Dumas 2005], ist die tabellenorientierte und formulargesteuerte Sichtweise auf Unternehmensabläufe immer noch Standard und macht die Extraktion von relevanten Daten für den jeweils betrachteten Prozess aus den beteiligten Transaktionssystemen und deren Beschreibung in Form von Prozessinstanzen notwendig. Dies erfordert spezifische Kenntnis über beteiligte IT-Systeme, meist Datenbanken. In Einzelfällen sind nachträgliche Ergänzungen (z.B. Zeitstempel) erforderlich, was die Datenerhebung zu einer komplexen Angelegenheit und gleichzeitig zur Achillessehne der Prozessanalyse macht.

3 Aufbau und Ziel der Prozessanalyse

Als Grundlage für das Fallbeispiel dienen die innerhalb eines SAP-Systems generierten Aufzeichnungen eines Standard-Bestellprozesses eines Industrie-Unternehmens (siehe Abb. 1). Die erhobenen Daten beziehen sich in erster Linie auf Bestellungen aus dem Verwaltungsbereich und umfassen neben regulären Warenbestellungen auch Buchungen von Schulungen zu Fachthemen und Bestellungen von Fachzeitschriften (siehe Abschnitt 4).

Abb. 1: Bestellprozess

Ausgehend von einer Bedarfsfeststellung wird dabei zunächst eine Bestellanforderung generiert. Nach erfolgter Freigabe wird der eigentliche Bestellvorgang initiiert und mit Waren- bzw. Rechnungseingängen abgeschlossen. Die Ausführung dieses Prozesses innerhalb des SAP-Systems resultiert in der Generierung von Daten in einer Vielzahl unterschiedlicher Tabellen, die vor der Verarbeitung mittels Process Mining zunächst konsolidiert werden müssen. Von diesem Vorgehen wird an dieser Stelle bewusst abstrahiert. Um den Umfang der Datenbasis zu begrenzen, werden im Folgenden ausschließlich abgeschlossene Fälle innerhalb eines Geschäftsjahres betrachtet.
Für die Durchführung der Prozessanalyse wird auf die zwei etablierten Software-Komponenten ProM und Disco zurückgegriffen. Hinweise für die Durchführung der Einzelanalysen innerhalb

dieser Anwendungen werden am Ende der jeweiligen Abschnitte in eckigen Klammern angegeben.

ProM ist eine Open-Source Process Mining Software aus dem akademischen Bereich (*processmining.org*), die eine Vielzahl wissenschaftlicher Ansätze zur Prozessanalyse, vorrangig der Modellrekonstruktion, enthält. Aufgrund der Plugin-orientierten Architektur eignet sich ProM besonders für die Erprobung und Evaluation wissenschaftlicher Ansätze und ist deshalb nah am aktuellen Stand der Forschung. Die wesentliche Stärke von ProM in Bezug auf Prozessaudits ist die vorhandene Unterstützung von Konformitätschecks mittels mathematischer Verifikation. ProM wird an der TU Eindhoven unter Federführung von Wil van der Aalst entwickelt.
Disco ist eine Process Mining Software der Firma Fluxicon (*fluxicon.com*). Analog zu weiteren kommerziellen Produkten dieser Sparte konzentriert sich Disco darauf, Auditoren mithilfe von verschiedensten Filtern und Darstellungsmöglichkeiten von Prozessinformationen dabei zu unterstützen, sich ein detailliertes Bild des tatsächlichen Prozessablaufs zu machen und abnormes Verhalten zu identifizieren. Disco beinhaltet die wesentlichen Aspekte von ProM, arbeitet jedoch speziell im Hinblick auf große Datenbestände stabiler und weist zudem ein erheblich intuitiveres Handling auf.

Anforderungen an die Prozessanalyse
Im Rahmen der Prozessanalyse soll das Ist-Verhalten des vorgestellten Bestellprozesses analysiert und Abweichungen zum Soll-Verhalten herausgearbeitet werden. Viele Unternehmen verfügen noch immer über keine aussagekräftigen, faktenbasierten End-to-End Prozesskennzahlen, weshalb für sie die Analyse der Prozessleistung im Vordergrund steht. Die Information über die unterschiedliche Art und Weise wie Prozesse ausgeführt werden, ist wesentlich für die Revision, weil gerade in selten auftretenden Prozessvarianten Risiken und Auflagenverstöße liegen können.
Die Anforderungen für die Analyse des Bestellprozesses werden in die drei Bereiche „Prozesskomplexität“, „Prozessperformance“ und „Risiko und Sicherheit“ untergliedert. Bezüglich der Komplexität soll untersucht werden, welche Aktivitäten bei der Prozessausführung involviert sind und wie oft diese auftreten. Speziell Aktivitäten, mit denen ein bestimmtes Risiko verbunden ist, spielen hierbei eine Rolle (Hinweise für die Unterlassung von erforderlichen Bewilligungen). Zusätzlich soll festgestellt werden, welche vom geplanten Prozessverlauf abweichenden Ausführungsarten existieren und welche davon fehlerhaftes bzw. nichtkonformes Verhalten beschreiben. Im Hinblick auf die Prozessperformance soll neben der Durchlaufzeit des Gesamtprozesses auch die Ausführungsdauer einzelner Aktivitäten analysiert werden, um kritische Bereiche zu identifizieren, die den Gesamtprozess verlangsamen. Obwohl in beiden Bereichen bereits risikobehaftetes Verhalten im Sinne ungewöhnlicher Prozessvarianten oder auffälliger Ausführungszeiten analysiert wird, soll sich der Bereich Risiko und Sicherheit auf spezifische Sicherheitsanforderungen konzentrieren, die für den vorliegenden Prozess gelten. Hierbei wird speziell auf die Funktionstrennung (Separation of Duty) eingegangen [Botha 2001].

4 Analyse der Prozesskomplexität

Der erste Schritt der Prozessanalyse besteht in der Betrachtung verschiedener globaler Statistiken, um einen Eindruck von der Komplexität zu gewinnen. Im vorliegenden Fall enthält die Logdatei 35.309 Vorkommen von 8 verschiedenen Aktivitäten die insgesamt 9.236 Bestellvorgänge formen und von 239 verschiedenen Personen ausgeführt wurden. Abb. 2 zeigt

identifizierte Aktivitäten und ihre Häufigkeit. Bereits hier fällt auf, dass es zu Abweichungen vom geplanten Prozessverhalten gekommen ist und dass das idealisierte Prozessmodell der Praxis nicht gerecht wird. Obwohl jede Bestellanforderung tatsächlich zu einer Bestellung geführt hat, liegt nicht in jedem Fall eine Freigabe vor. Zudem liegt die Zahl an Wareneingängen deutlich unter der Zahl von Bestellungen. Hohe Anteile vermeintlich fehlerhafter Abweichungen (hier ca. 30%) lassen sich jedoch oftmals durch legitime Abläufe erklären, die nicht im Prozessmodell berücksichtigt sind. In diesem Fall sind dies vorrangig Abonnements und Schulungen, die im System zwar als Bestellung hinterlegt, jedoch nicht mit einem Wareneingang abgeschlossen wurden. Die Abweichung hinsichtlich zusätzlicher Aktivitäten wird als unkritisch eingestuft, weil es sich dabei um übliche Aktivitäten in Zusammenhang mit Bestellprozessen handelt. Während alle Prozessinstanzen plangemäß mit einer Bestellanforderung beginnen, zeigt Abb. 2. eine verhältnismäßig große Zahl von Endaktivitäten. Hierbei bedürfen die Fälle „Bestellung anlegen“ und „Anzahlung“ einer genaueren Betrachtung, um deren Ursprung zu klären.

Aktivität	Häufigkeit	Endaktivität
Rechnungseingang	10974	ja
Bestellanforderung	9236	nein
Bestellung anlegen	9236	ja
Freigabe	9205	nein
Wareneingang	6477	ja
Nachbelastung	101	ja
Anzahlung	58	ja
Kontenpflege	22	ja

Abb. 2: Aktivitätsstatistik

Bezüglich der Komplexität von Prozessinstanzen liegt laut Abb. 3 die Anzahl der Aktivitäten in den meisten Fällen bei 4, jedoch gibt es einzelne Ausreißer mit ungewöhnlich vielen Aktivitäten.

	Minimal	Maximal	Schnitt
Aktivitäten pro Instanz	4	54	4
Aktivitätsklassen pro Instanz	4	6	4

Abb. 3: Instanzenstatistik

Auch die Zahl unterschiedlicher Aktivitäten (Klassen) zeigt, dass offensichtlich viele „zu kurze“ Fälle existieren. In der folgenden Variantenanalyse muss überprüft werden, ob es plausible Gründe für das Fehlen von Aktivitäten gibt (*[**Disco**: Statistics - Overview] [**ProM**: Log Summary]).*

Variantenanalyse

Im Rahmen der Variantenanalyse soll festgestellt werden, auf welche unterschiedlichen Arten der Bestellprozess (bzgl. der Reihenfolge von Aktivitäten) ausgeführt wird und inwiefern diese Varianten vom geplanten Prozessverlauf abweichen. Die Erstellung eines Modells, das die

verschiedenen Aktivitätsübergänge veranschaulicht, eignet sich gut um die Gesamtvarianz des Prozesses einschätzen zu können. Eingeschränkt auf die Hauptaktivitäten des Bestellprozesses zeigt Abb. 4 die zugehörigen Übergänge gemäß dem Fuzzy-Mining Ansatz [Günther 2007].

Aktivitätssequenz	Häufigkeit	
Bestellanforderung - Freigabe - Bestellung anlegen - Wareneingang - Rechnungseingang	5463	59,15%
Bestellanforderung - Freigabe - Bestellung anlegen - Rechnungseingang	2727	29,53%
Bestellanforderung - Bestellung anlegen - Freigabe - Wareneingang - Rechnungseingang	300	3,25%

Abb. 4: Modell zu Aktivitätsübergängen

Hier steht nicht die Struktur des Prozesses im Vordergrund, sondern die verschiedenen Übergänge zwischen jeweils zwei Aktivitäten und ihre Häufigkeit. Werte innerhalb von Aktivitäten beschreiben deren Häufigkeit bzgl. Prozessinstanzen. Generell ist bei dieser Form der Modellinterpretation zu beachten, dass aufgrund evtl. existierender UND/ODER Verknüpfungen im zugrunde liegenden Ausführungsmodell Übergänge sich gegenseitig bedingen oder ausschließen können. Deshalb ist die Analyse von Einzelübergängen nicht immer zielführend und macht weitere Untersuchungen erforderlich, um nichtkonformes Verhalten zweifelsfrei zu bestimmen. Im vorliegenden Fall wird offensichtlich entgegen dem geplanten Verhalten in insgesamt 367 Fällen eine Bestellung angelegt, bevor eine Freigabe vorliegt. Es bleibt zu klären, in welchen Fällen diese nachträglich erfolgt oder ausbleibt. Auffällig ist auch, dass in 317 Fällen direkt nach der Freigabe Wareneingänge verzeichnet werden. Deshalb ist es notwendig, mit weiteren Mitteln auffällige Instanzen zu extrahieren und genauer zu untersuchen (*[**Disco**: Map – Frequency Modus] [**ProM**: Mining – Fuzzy Miner])*

Die Gesamtzahl unterschiedlicher Prozessvarianten liegt bei 97. Obwohl dies für den strukturell relativ einfachen Prozess zunächst viel scheint, wird bei näherer Betrachtung klar, dass die meisten Varianten aus der Parallelität und dem Mehrfachauftreten von Waren- und Rechnungseingängen resultieren. Abb. 5 zeigt die drei am häufigsten auftretenden Prozessvarianten.

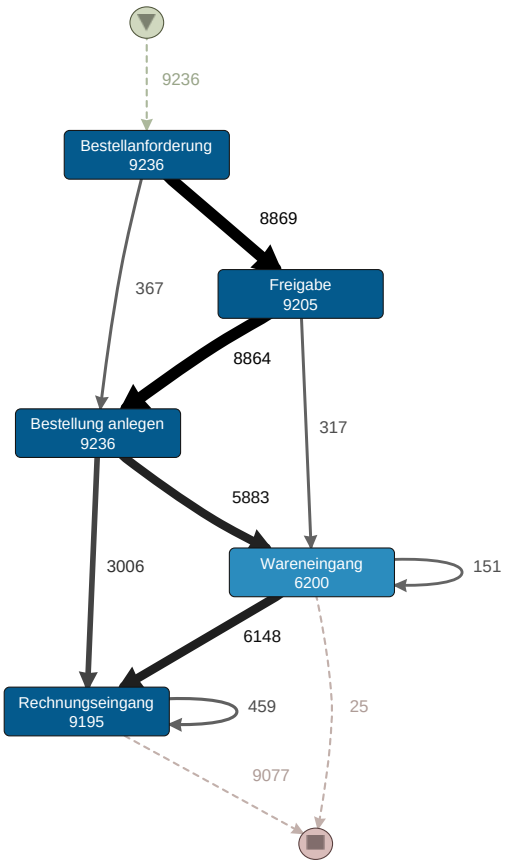

Abb. 5: Häufigste Prozessvarianten

In 59,15% der Fälle wird der Prozess exakt gemäß dem Prozessmodell ausgeführt. Deutlich erkennbar ist auch der relativ große Anteil von Varianten, die ohne Wareneingang ablaufen. Obwohl diese zum Großteil als korrekt bewertet werden dürften, kann dieses Verhalten auch eine ungewünschte Abweichung darstellen. Deshalb muss hier durch weitere Analysen und ggf. Interviews mit Prozessbeteiligten Ursachenforschung betrieben werden. Bei der dritthäufigsten Variante muss überprüft werden, ob nachträgliche Freigaben zulässig waren (*[**Disco**: Cases] [**ProM**: Mining – Log Summary, Exports – Grouped MXML Log (sequences) – Log Summary])*

Anstatt jede einzelne Variante im Detail zu untersuchen, werden diejenigen Instanzen extrahiert, die bezüglich der vorgegebenen Analysezielen relevante Eigenschaften aufweisen. Die ProM-Plugins *LTL-Checker* und *SCIFF-Checker* erlauben die detaillierte Auswahl von Instanzen mit spezifischen Eigenschaften anhand von Formeln temporaler Logik. So können beispielsweise Instanzen isoliert werden, die eine Aktivität A nach einer Aktivität B enthalten oder bei denen eine bestimmte Person eine bestimmte Aktivität ausführt. Disco stellt für diesen Zweck diverse Filter

zur Verfügung. Um nicht den Rahmen des Beitrags zu sprengen, wird im Folgenden nur auf die Prüfung von Bestellungen vor/ohne Freigabe eingegangen.

Die Filterung nach Bestellungen, die vor der eigentlichen Freigabe erfolgt sind, resultiert in 336 Prozessinstanzen. Aufgrund dieser verhältnismäßig hohen Zahl liegt der Schluss nahe, dass eine interne Bearbeitungsrichtlinie oder „common practice“ existiert, die dieses Verhalten erklärt. Durch Rücksprachen mit den Prozessteilnehmern und Bereichsverantwortlichen muss die generelle Plausibilität dieser Abweichung oder die Rechtmäßigkeit im Einzelfall geklärt werden. Das Vorliegen von Informationen über den konkreten Ablauf infrage stehender Prozessausführungen und weitere Parameter wie z.B. beteiligte Personen ist hierbei hilfreich.
Dies trifft in besonderem Maße auf die isolierten 31 Instanzen zu, die ohne Freigabe durchgeführt wurden. Ein Blick auf die Ausführungsdauer (die in Betrugsfällen oft außergewöhnlich kurz ausfällt) lässt hier keinen Schluss auf Prozessmanipulation zu. Die Tatsache, dass keine Person mehr als eine Aktivität durchführt, deutet eher auf unbeabsichtigtes Fehlverhalten hin. Dennoch kann sich ein Blick auf die beteiligten Akteure lohnen, um evtl. kooperatives Vorgehen zu Betrugszwecken oder fehlerhafte Prozessbearbeitung von Personengruppen aufzudecken. Erste Hinweise kann hierbei die relative Häufigkeit der Ausführung von Aktivitäten einer Person liefern. Innerhalb der 31 fraglichen Instanzen wird die Bestellanforderung in 83% von Benutzer 205 und das Anlegen der Bestellung in 77% von Benutzer 436 durchgeführt. Der Wareneingang wird vorrangig von Benutzer 206 (49%) und Benutzer 167 (37%) verbucht, der Rechnungseingang von Benutzer 324 (50%) und Benutzer 388 (29%). Diese Information lässt sich zwar nicht direkt ablesen, ist aber mit überschaubarem Aufwand manuell berechenbar. Selbst wenn ein Vergleich mit den relativen Häufigkeiten innerhalb aller vorliegenden Prozessinstanzen aufgrund der relativ kleinen Teilmenge von 31 Instanzen wenig vielversprechend ist, können zumindest die Hauptverursacher bestimmt werden.

Eine weitere Möglichkeit im Zusammenhang mit kooperativem Verhalten ist die Betrachtung der Beziehungen von Personen innerhalb von Prozessinstanzen. Die Extraktion eines sog. Social Networks erlaubt beispielsweise die Betrachtung des „work handover“, also wer wem den nächsten Prozessschritt „übergibt“. Abb. 6 zeigt das resultierende Netzwerk für die obigen Benutzer zusammen mit den Übergabeanteilen (*[**ProM**: Analysis – LTL Checker, Analysis – Social Network Miner])*

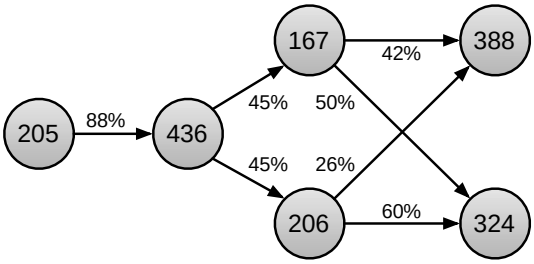

Abb. 6: Soziales Netzwerk

5 Analyse der Prozessperformance

Obwohl bei der Prozessauditierung Analysen bezüglich der Prozesskomplexität und insbesondere dem Variantenreichtum von Prozessen im Vordergrund stehen, kann im Rahmen von Performanceanalysen ein Augenmerk auf Instanzen mit ungewöhnlich langer/kurzer Ausführungszeit gelegt werden. Sofern die Verbesserung der Leistungsfähigkeit des Prozesses Ziel des Audits ist, eignet sich eine Performanceanalyse dazu, Information für die Ableitung von Optimierungsmaßnahmen zu generieren. Dabei stehen die Durchlaufzeit sowohl des Gesamt- als auch spezifischer Teilprozesse und die Bearbeitungsdauer einzelner Aktivitäten im Vordergrund.
Die Gesamtperformance des Prozesses wird anhand eines Graphen abgeschätzt, der die mittlere Übergabezeit zwischen den Hauptaktivitäten des Prozesses darstellt (siehe Abb. 7).

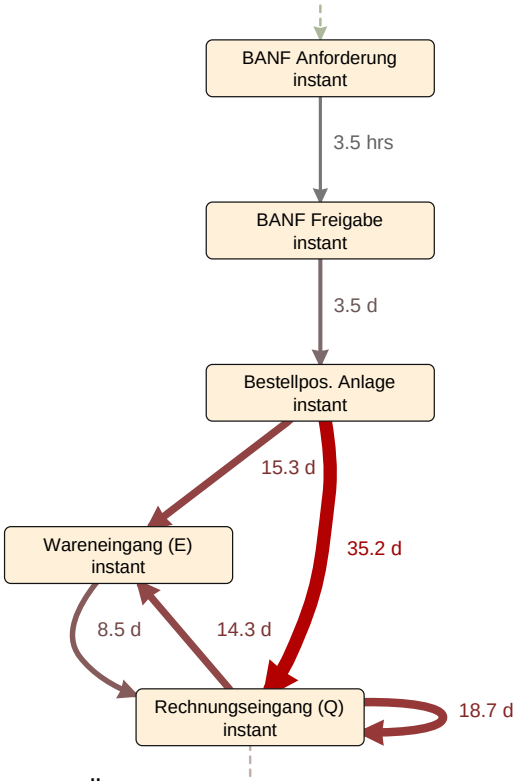

Abb. 7: Übergabezeiten zwischen Prozessaktivitäten

Während Freigaben im Mittel 3,5 Stunden nach der Bestellanforderung erfolgen, vergehen im Schnitt 15 Tage bis zum ersten Waren- und 35 Tage bis zum ersten Rechnungseingang. Diese Werte müssen vor dem Hintergrund der üblichen Geschäftstätigkeit und ggf. existierenden Liefervereinbarungen interpretiert werden.

Die Durchlaufzeit von Bestellungen von Bestellanforderung bis zum letzten Waren- bzw. Rechnungseingang kann mittels diverser Statistiken innerhalb der Softwarekomponenten nachvollzogen werden und liegt zwischen 14 Stunden und einem Jahr und 6 Tagen. Die meisten Bestellungen (90%) dauern maximal 75 Tage. Um einen genaueren Blick auf unnormal lange/kurze Bestellungen zu werfen, können bei Disco entsprechende Prozessinstanzen per Filter ausgewählt werden. Im vorliegenden Fall sind sehr lange Laufzeiten auf Bestellungen zurück zu

führen, die für längerfristige Vertragsbindungen in Kombination mit Monatsabrechnungen stehen und deshalb mehrfache, mehr oder minder regelmäßige Rechnungseingänge aufweisen. Bei auffällig kurzen Bearbeitungszeiten kann eine detaillierte Betrachtung der betreffenden Instanzen unter besonderer Berücksichtigung beteiligter Personen analog zu Abschnitt 4.1 sinnvoll sein.

Bei Disco ist die Information über die Dauer einzelner Cases und die grafische Ansicht für Aktivitätsübergänge innerhalb von Cases direkt verfügbar. Auch die Anzahl aktiver Instanzen und Aktivitäten pro Tag kann nachvollzogen werden. Bei der Verwendung von ProM ist zu beachten, dass in manchen Fällen nicht nur das Datum und die Uhrzeit der Fertigstellung einer Aktivität, sondern auch deren Beginn benötigt wird. Dies muss, falls möglich, schon bei der Datenextraktion berücksichtigt oder nachträglich annotiert werden (*[**Disco**: Filter – Attribute – Activity, Map – Performance (Mean), Cases – Statistics], [**ProM**: Filter – Eventfilter, Analysis – Perf. Sequence Analysis, Basic Perf. Analysis]).*

6 Analyse von Risiko und Sicherheit

Auch wenn im Rahmen der Komplexitäts- und Performanceanalysen schon Risikoabschätzungen mithilfe der Identifikation und Beurteilung von Verhaltensauffälligkeiten durchgeführt werden können, gibt es sicherheitsorientierte Prozesseigenschaften in Form systemspezifischer Benutzerbefugnisse, die über den in Isolation betrachteten Prozess hinausgehen und einer separaten Überprüfung bedürfen. Ein starker Fokus liegt hierbei auf rollenbasierter Zugangskontrolle (RBAC) und Vorgaben zur Funktionstrennung. Solche Richtlinien sind in vielen Fällen mithilfe mathematischer Verifikation überprüfbar. Existierende wissenschaftliche Ansätze bieten hierfür die Möglichkeit, Prozessanforderungen in Linearer Temporaler Logik (LTL) zu formulieren und anschließend auf Prozesslogs zu verifizieren [van der Aalst 2005]. Im Wesentlichen werden dabei Prozessanforderungen in Form mathematischer Formeln beschrieben. LTL bietet in diesem Zusammenhang die Möglichkeit, zeitliche Abfolgen zu definieren (z.B. „Wenn A ausgeführt wird, dann immer auch B“). Die Verifikation klärt anschließend die Frage, ob einzelne Prozessinstanzen diese Formeln erfüllen. Für den vorliegenden Bestellprozess gelten folgende Anforderungen bzgl. der Funktionstrennung, die Prozessmissbrauch zu kriminellen Zwecken verhindern sollen:

1. **Personen, die Bestellanforderungen erstellen, dürfen diese nicht selbst freigeben.** Dieses 4-Augen-Prinzip soll sicherstellen, dass die Freigabe dadurch umgangen wird, dass Personen ihre eigenen Bestellanforderungen freigeben.
2. **Personen, die Bestellungen anlegen, dürfen keine Wareneingänge bestätigen.** Durch diese Anforderung soll erschwert werden, dass Personen nicht-reale Bestellvorgänge über gefälschte Lieferantenkonten initiieren.

Um diese Anforderungen zu überprüfen, wird für Prozessinstanzen mit entsprechenden Aktivitäten und dem in ProM verfügbaren LTL-Preset „*Exists person doing task A and B“* überprüft, ob die Funktionstrennung verletzt wurde. Während im ersten Fall keine Verletzungen vorliegen, beinhalten die betreffenden 6.200 Bestellvorgänge für Anforderung 2 tatsächlich drei Verletzungen, einmal durch Benutzer 293 und zweimal durch Benutzer 269. In allen drei Fällen wurde die Bestellung von Benutzer 20 freigegeben. Abb. 8 zeigt, dass die Bestellanforderung von unterschiedlichen, Rechnungseingänge jedoch von derselben Person verbucht wurden. Ob in diesen Fällen tatsächlich eine Verletzung vorliegt, muss im Einzelfall geklärt werden.
*[**ProM**: Analysis – LTL Checker]*

Anforderung	Freigabe	Bestellung anlegen	Wareneingang	Rechnungseingang
267	20	269	269	388
77	20	269	269	388

Abb. 8: Verletzungen der Funktionstrennungs-Anforderungen.

Für die Überprüfung von Autorisierungs-Richtlinien besteht die Möglichkeit, eine Matrix zu generieren, die widergibt, welche Prozessaktivitäten von einzelnen Personen durchgeführt werden. Diese Matrix manuell zu analysieren ist sehr aufwändig und in den meisten praktischen Fällen nicht in akzeptabler Zeit durchführbar. Die aktuelle Version von ProM enthält deshalb ein Plugin, welches die direkte Überprüfung von RBAC-Richtlinien auf Logdaten erlaubt, sofern ein Organisationsmodell bereitgestellt wird, welches Benutzer und deren Rollenzugehörigkeit beschreibt [Baumgraß 2011]. Berechtigungen werden auf Aktivitätsebene spezifiziert und intern auf LTL-Formeln zurückgeführt, die dann mithilfe existierender Model-Checking Mechanismen verifiziert werde (*[**ProM**: Analysis: Originator by Task Matrix]).*

7 Erfahrungen mit der Einsatzfähigkeit von Process Mining

Die Voraussetzung für die Anwendung von Process Mining ist die Existenz eines Prozesslogs der Systemaktivitäten gemäß fest definierter Aktivitäten und gruppiert nach Prozessinstanzen enthält. Das grundsätzliche Problem hierbei ist die Diskrepanz zwischen der wirtschaftlichen und technischen Sicht auf Geschäftsprozesse, die sich darin äußert, dass Prozessaktivitäten, wenn auch in Modellen hinterlegt, nicht in derselben Form in den beteiligten Informationssystemen implementiert sind. Ausgehend von Referenz-Prozessmodellen müssen deshalb zunächst auf Basis von Systemereignissen Prozessaktivitäten identifiziert und eine Strategie für deren korrekte Extraktion gefunden werden.
Für die Durchführung von computergestützten Prozessanalysen auf Basis von Process Mining ist fachliche Expertise bzgl. Datenbank- und Informationssystemen sowie in Einzelfällen informatisches Grundlagenwissen (formale Modelle, mathematische Verifikation) notwendig. Für Unternehmen bedeutet dies, dass notwendiges Know-How entweder durch geeignete Mitarbeiterschulungen oder das Einwerben von Fachpersonal generiert, oder durch die Inanspruchnahme von Leistungen externer Anbieter erbracht werden muss.
Die korrekte Interpretation von Prüfungsergebnissen auf Basis von Process Mining erfordert in jedem Fall eine klare und die Prüfung in allen Teilbereichen begleitende Kommunikation zwischen Prüfer und Prozessverursacher/-verantwortlichen, um aussagekräftige Ergebnisse zu produzieren.
Obwohl durch die Verwendung existierender Softwarekomponenten die Hauptinteressen von Prozessprüfungen grundsätzlich abgedeckt sind, lässt sich in vielen Fällen das Analyseziel nur durch zusätzlichen manuellen Aufwand oder die Kombination verschiedener Komponenten erreichen. Es ist jedoch absehbar, dass existierende Produkte ihren Funktionsumfang kontinuierlich erweitern werden, um dem wissenschaftlichen Stand der Technik näher zu kommen.
Im Vergleich zu manuellen Stichprobenverfahren berücksichtigen die verwendeten Methoden die Gesamtheit verfügbarer Prozessdaten und generieren damit Ergebnisse, die sich aufgrund ihrer höheren Aussagekraft besser als Entscheidungsgrundlage für Konformitätsfragen eignen.
Generell kann durch den Einsatz von Process Mining bei der Prüfung von Geschäftsprozessen deshalb ein detailliertes Verständnis für die reale Prozessausführung gewonnen werden. Der erreichbare Mehrwert hängt dabei wesentlich von der korrekten Datenextraktion und -aufbereitung

ab. Die Entwicklung prozessorientierter Informationssysteme für die Abwicklung betrieblicher Abläufe auf Basis klar definierter Prozessmodelle (PAIS) dürfte in diesem Zusammenhang die Anwendbarkeit von Process Mining in Zukunft weiter erhöhen.

8 Literatur

[van der Aalst 2011] *van der Aalst W.:* Process Mining – Discovery, Conformance and Enhancement of Business Processes. Springer, 2011.

[van der Aalst 2010] *van der Aalst W., van Hee K., van der Werf J., Verdonk M.:* Auditing 2.0: Using Process Mining to Support Tomorrow's Auditor. IEEE Computer 43(3):90-93, 2010.

[van der Aalst 2005] *van der Aalst W., de Beer H., van Dongen B.:* Process Mining and Verification of Properties: An Approach Based on Temporal Logic. OTM Conferences, S. 130-147, 2005.

[Accorsi 2011] *Accorsi, R.*: Rafael Accorsi: Business Process as a Service: Chances for Remote Auditing. IEEE Computer Software and Applications Conference, S. 398-403, 2011.

[Accorsi 2012] *Accorsi R., Stocker T.:* On the Exploitation of Process Mining for Security Audits: The Conformance Checking Case. ACM Symposium on Applied Computing, S. 1709-1716, 2012.

[Baumgraß 2011] *Baumgraß A., Baier T., Mendling J., Strembeck M.:* Conformance Checking of RBAC Policies in Process-Aware Information Systems. Business Process Management Workshops, S. 435-446, 2011.

[Botha 2001] *Botha R., Eloff. J.:* Separation of Duties for Access Control Enforcement in Workflow Environments. IBM Systems Journal 40(3):666-682, 2001.

[Carlin 2007] *Carlin A., Gallegos F.:* IT Audit: A Critical Business Process. IEEE Computer 40(7):87-89, 2007.

[Cook 1999] *Cook J., Wolf A.:* Software Process Validation: Quantitatively Measuring the Correspondence of a Process to a Model. ACM Transactions on Software Engineering Methodology 8(2):147 - 176, 1999.

[van Dongen 2009] *van Dongen B., Alves de Medeiros A., Wen L.:* Process Mining: Overview and Outlook of Petri Net Discovery Algorithms. Transactions on Petri Nets and Other Models of Concurrency II, S. 225-242, 2009.

[Dumas 2005] *Dumas M., van der Aalst W., ter Hofstede A.:* Process Aware Information Systems: Bridging People and Software Through Process Technology. Wiley-Interscience, 2005.

[Günther 2007] *Günther C., van der Aalst W.:* Fuzzy Mining: Adaptive Process Simplification Based on Multi-Perspective Metrics. Business Process Management (BPM'07), S. 328-343, Springer, 2007.

[Lowis 2011] *Lowis L., Accorsi, R.:* Vulnerability Analysis in SOA-Based Business Processes. IEEE Trans. on Services Computing 4(3):230-242, 2011.