

Process Mining: Op zoek naar het bekende onbekende

"Data is de nieuwe olie" is een uitspraak die we de laatste jaren steeds vaker horen. Data is namelijk de grondstof die het informatietijdperk mogelijk maakt. Op dezelfde manier was olie de grondstof die het industriële tijdperk mogelijk maakte. In tegenstelling tot olie raakt data echter nooit op. In tegendeel, de hoeveelheid data die we produceren neemt jaar na jaar toe. In het informatietijdperk is het daarom onze uitdaging om waarde uit deze grote hoeveelheden data te halen.

door Joris Mens en Remco Bunder

Het bekende onbekende

Process mining helpt ons hierbij. Process mining is een techniek waarmee we een visuele reconstructie van een proces maken met daadwerkelijke procesgegevens als input. Dit doen we door een procesmodel te reconstrueren, uitgaande van transactiegegevens uit informatiesystemen. Deze gegevens noemen we ook wel een event log. In een event log legt een systeem de statussen vast waar een proces doorheen loopt. Denk bijvoorbeeld aan het afhandelen van een serviceaanvraag of het plaatsen van een inkooporder. Op basis van deze statussen kan process mining-software een procesmodel maken, waarmee het daadwerkelijke proces visueel inzichtelijk wordt. Vervolgens kunnen we aan dit procesmodel performance-informatie toevoegen, die bijvoorbeeld laat zien hoe lang bepaalde stappen duren of welke paden in het proces het vaakst gebruikt worden.

Figuur 1, 2, en 3 op de pagina hiernaast geven weer hoe een procesmodel gereconstrueerd wordt op basis van een event log en hoe hierop performance-informatie wordt weergegeven

We kunnen process mining op brede schaal toepassen, omdat er steeds meer systemen zijn die transacties in detail bijhouden. Bovendien wordt de behoefte naar process mining groter omdat we onze complexe bedrijfsprocessen snel willen veranderen en bijsturen op basis van realtime informatie. Vaak weten we dat een proces niet optimaal verloopt maar hebben we onvoldoende inzicht in de precieze oorzaak. Met process mining gaan we dus op zoek naar het bekende onbekende. In dit artikel beschrijven we hoe process mining ons heeft geholpen om waarde te scheppen uit procesdata. We verkennen eerst de principes achter process mining. Vervolgens vertellen we hoe process mining een aantal processen binnen NS heeft helpen verbeteren, waaronder het oplossen van storingsmeldingen op stations.

Historie

Process Mining is de afgelopen tien jaar in opkomst geraakt. Begin 2000 is onder leiding van professor Wil van der Aalst aan de TU Eindhoven begonnen met de geautomatiseerde analyse van event logs. Hiermee zouden we kunnen stellen dat professor van der Aalst de grondlegger is van het fenomeen process mining. De oorsprong van Process Mining ligt echter nog veel verder terug in de tijd. Al in 1960 werden statussen van machines in kaart gebracht in een handgeschreven journaal. Software maakt het eenvoudiger om transacties automatisch bij te houden in een event log. Daarom zijn de mogelijkheden om process mining toe te passen gegroeid. Inmiddels bestaat er een groot aantal leveranciers van process mining software en zijn er verschillende universiteiten die op dit gebied onderzoek verrichten.

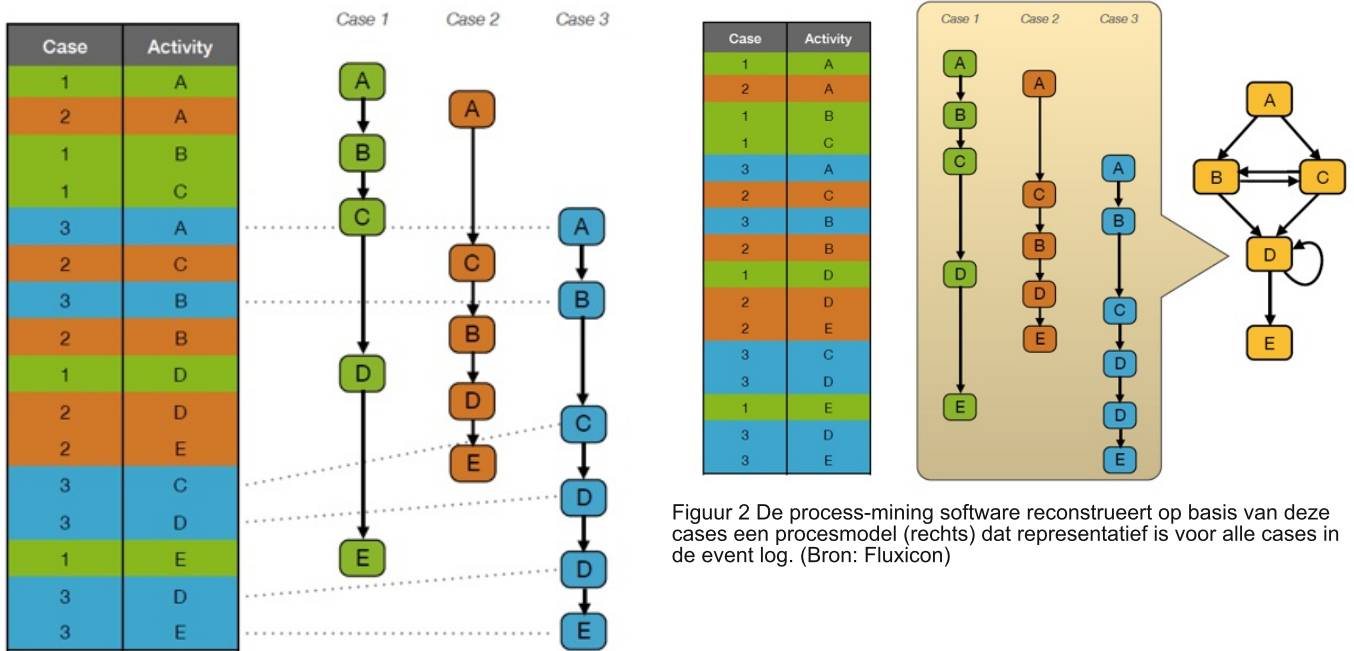
Process Mining kan gezien worden als subdiscipline binnen de bredere disciplines van Data Analytics en Business Intelligence. Hierbij wordt wel eens de vraag gesteld: Wat is het verschil tussen process mining en data mining? In tegenstelling tot data mining kijkt process mining naar de volgorde van events. De tijdstippen waarop transacties plaatsvinden spelen een centrale rol bij process mining en worden gebruikt om de prestaties van processen vast te stellen.

Daarnaast draagt process mining bij aan de LEAN-filosofie. Het is een ideaal middel om verspilling in processen inzichtelijk te maken en te bespreken met domeinexperts. Zo krijgen we bijvoorbeeld inzicht in bottlenecks in de vorm van wachttijden, stappen die worden herhaald, of procespaden waarbij wordt afgeweken van de norm.

De Business Process Lifecycle

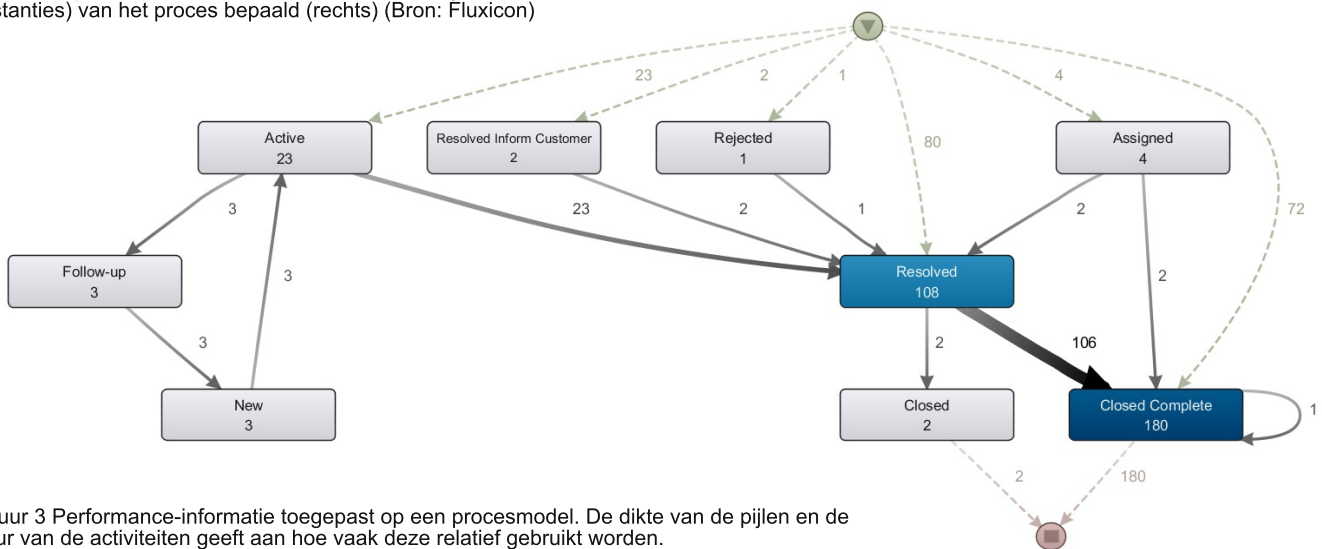
We definiëren een bedrijfsproces als een verzameling van activiteiten die waarde creëren. Het is van belang dat een bedrijfsproces aansluit op de bedrijfsstrategie

Process Mining bij NS



Figuur 1 Een event log (links) bevat een case ID en benoemt de activiteiten per case. Hiermee worden de afzonderlijke cases (instanties) van het proces bepaald (rechts) (Bron: Fluxicon)

Figuur 2 De process-mining software reconstrueert op basis van deze cases een procesmodel (rechts) dat representatief is voor alle cases in de event log. (Bron: Fluxicon)



Figuur 3 Performance-informatie toegepast op een procesmodel. De dikte van de pijlen en de kleur van de activiteiten geeft aan hoe vaak deze relatief gebruikt worden.

(alignment) en dat het zo efficiënt mogelijk verloopt. Een proces staat niet voor altijd vast, aangezien bedrijven zich aanpassen aan een dynamische omgeving. We spreken daarom ook wel over de process lifecycle. Dit is een cyclus van continue verbetering, waarin we een proces monitoren en evalueren om vervolgens tot een verbeterde versie van het proces te komen. We schetsen hier de process lifecycle om duidelijk te maken op welke vlakken process mining bijdraagt aan continue verbetering.

De process lifecycle definiëren we als volgt:

1. Ontwerp;
2. Implementatie;
3. Monitoring;
4. Evaluatie.

De process lifecycle begint bij (1) het ontwerpen van een proces op basis van een aantal requirements.

Vervolgens richten we bij (2) de implementatie een systeem in dat het proces uitvoert en ondersteunt. Vervolgens (3) monitoren we in dit systeem de lopende processen. Wanneer we procesgegevens over een langere periode hebben verzameld kunnen we (4) een evaluatie uitvoeren. Deze evaluatie geeft input voor een (her)ontwerp van het proces, waarmee we een nieuwe iteratie van de process lifecycle bereiken.

Process mining draagt bij aan de fases Monitoring en Evaluatie. We kunnen de prestaties van het lopende proces meten en eventueel voorspellingen doen over de mogelijke uitkomsten of resterende doorlooptijd van het proces. Ook kunnen we een event log met gegevens van bijvoorbeeld het afgelopen half jaar evalueren en uitspraken doen over de prestaties van het proces. Deze analyses bieden waardevolle input om vast te stellen waar in het proces mogelijke verspilling plaatsvindt.

Storingsmeldingen

Met behulp van process mining krijgt de informatieanalist een gedetailleerd inzicht in de IST-situatie (op basis van daadwerkelijke gegevens) en kan daarmee beschrijven waar het proces afwijkt van de SOLL-situatie. De informatieanalist kan vervolgens de requirements opstellen voor een verbeterde iteratie van het proces.

Aan de slag met process mining

In dit artikel hebben we een aantal keren de term event log genoemd. Een event log moet bestaan uit de volgende gegevens om geschikt te zijn voor process mining:

1. Een case ID (een unieke identifier voor een enkele instantie van het proces)
2. Een activiteit
3. Een timestamp

Deze gegevens zijn voldoende om een visuele reconstructie te maken van het proces. Hiernaast kunnen er talloze andere gegevens worden toegevoegd die relevant zijn voor de analyse. Bij de NS bestaat er een proces voor het oplossen van storingsmeldingen op stations. Hierbij gaat het bijvoorbeeld om liften of roltrappen die buiten werking zijn, maar ook zaken als defecte verlichting op de perrons en kapotte koffieautomaten in de stationswinkels. Hierbij kunnen we bijvoorbeeld gegevens toevoegen over de locatie (stationsnaam), het objecttype (lift, roltrap, koffieautomaat, etc.). Dit geeft ons de mogelijkheid om op basis van deze gegevens te filteren en te vergelijken. Zo kunnen we bijvoorbeeld zien of een bepaald object meer storingsmeldingen heeft dan een ander object, en hoe snel deze storingsmeldingen opgelost worden. Zo krijgen we inzicht in mogelijke terugkerende storingsmeldingen op stationsobjecten.

Nadat we hebben geïdentificeerd in welk systeem de gegevens rondom dit proces worden bijgehouden, maken we een extract van de database met daarin de voor ons benodigde gegevens. Het daadwerkelijke uitvoeren van process mining gebeurt in een tool. Inmiddels zijn er verschillende tools op de markt. Binnen het wetenschappelijke domein gebruikt men vooral

ProM, een open-source pakket dat werd ontwikkeld bij de TU Eindhoven. Een andere populaire tool is Fluxicon Disco. Dit is een laagdrempelige process mining suite die eveneens voortkomt uit de inspanningen van onderzoekers van de TU/e. Naast de voornoemde tools die zich richten op een brede markt, bestaan er ook process mining tools waaronder Celonis en ProcessGold die toegespitst zijn op implementaties op specifieke systemen waaronder SAP of Salesforce. In onze praktijkcases maken we gebruik van Disco.

Het verkrijgen van de juiste data die geschikt zijn voor process mining kan een grotere uitdaging zijn dan je in eerste instantie zou verwachten. Dit heeft voornamelijk te maken met de datakwaliteit. Omdat we met process mining een afspiegeling willen maken van de realiteit, is het belangrijk dat de procesgegevens deze accuraat representeren. Wordt een opgeloste storing bijvoorbeeld altijd direct afgemeld, of worden deze pas aan het eind van de dag in het systeem geboekt? Een ander mogelijk obstakel is dat procesgegevens vaak over meerdere systemen verspreid zijn. Afzonderlijke systemen hanteren bijvoorbeeld verschillende identifiers of activiteitsnamen voor wat feitelijk hetzelfde is. Voordat de daadwerkelijke analyse plaatsvindt zit de grootste uitdaging in process mining vaak in het opsporen en opschonen van de data.

Een ander belangrijk aspect bij het uitvoeren van process mining is het betrekken van domeinexperts uit de organisatie. Het proces dat je gaat analyseren wordt door verschillende personen in de organisatie uitgevoerd.

Praktijkcase: Storingsmeldingen op stations

In onze praktijkcase onderzoeken we het eerder genoemde proces van storingsmeldingen op stations. Onze afdeling (NS Stations) coördineert dit proces. Dit gaat van het aannemen van storingsmeldingen van de mensen die op het station werken, het aansturen van aannemers om de storing op te lossen, tot de administratieve en financiële afhandeling. Storingsmeldingen worden op verschillende manieren ontvangen. Eerst ging dit voornamelijk via telefoon of e-



Joris Mens

Joris Mens (1989) studeerde Business Informatics aan de Universiteit Utrecht en specialiseerde zich in Business Process Management. Zijn afstudeeronderzoek naar het toepassen van procesmanagement voor het borgen van kwaliteitsnormen in de zorg werd bekroond met de BPM Student Award.

Joris werkte daarnaast als onderzoeker en docent op de Hogeschool Utrecht. Hij gaf colleges en workshops over BPM en aanverwante onderwerpen, zoals business rules, business-IT alignment, en process mining.

Tegenwoordig ontwikkelt hij IT-strategie bij de Nederlandse Spoorwegen en past process mining toe om processen te verbeteren.

Process Mining bij NS

mail. Een medewerker van de servicedesk stations verwerkt deze binnenkomende meldingen in het facilitaire systeem. Tegenwoordig bestaat er ook een interne app waarmee medewerkers een storingsmelding kunnen plaatsen. De gebruiker van de app geeft aan om wat voor soort object het gaat en op welke locatie dit zich bevindt. Daardoor kan de storingsmelding direct in het facilitaire systeem worden geboekt zonder tussenkomst van een servicedeskmedewerker. Daarnaast zijn er stationsobjecten (waaronder liften) die door middel van een webservice automatisch een storing kunnen melden. In totaal hebben we dus vier manieren van het melden van een storing: telefoon, e-mail, app, en webservice.

In onze praktijkcase wilden we onderzoeken of de manier van het melden van de storing invloed heeft op de doorlooptijd van het proces. We wilden bijvoorbeeld te weten komen of de moderne manieren van storingsmeldingen (app en webservice) een snelheidswinst opleveren ten opzichte van meldingen die per e-mail of telefoon worden gedaan.

Meldingswijze	Gemiddelde doorlooptijd	Aantal processtappen
Telefoon	13.3 uur	18
App	22 min	8
Mail	15.6 uur	7
Webservice	2.6 uur	7

Uit de resultaten blijkt dat het melden van een storing via de app of webservice inderdaad snelheidswinst oplevert. Toch duurt het afhandelen van een melding via de webservice langer dan verwacht. Na enig onderzoek bleek dit veroorzaakt te worden door het feit dat meldingen via de webservice in een mailbox terechtkomen, die door een persoon worden uitgelezen. Deze meldingen komen dus niet direct terecht in het facilitaire systeem, zoals dat bij de app wel het geval is. Een uitdaging bij het invoeren voor de app is vooralsnog het filteren van dubbele meldingen. Waar er bij de methodes telefoon en mail een servicedeskmedewerker is die de melding aanneemt en kan koppelen aan

eventuele bestaande meldingen, is dit bij de app niet het geval. We zien echter dat het gebruik van de app het melden van storingsmeldingen toegankelijker maakt voor medewerkers en het de afhandeltijd verkort. Reden om deze methode van melden te stimuleren binnen de organisatie.

Naast het analyseren van de verschillen in de afhandeltijden tussen de verschillende methodes hebben we ook gekeken naar de eigenschappen van het proces. Het proces voor het afhandelen van storingsmeldingen was in de organisatie gedocumenteerd in de vorm van een procesmodel. Onze analyse toonde aan dat er in het proces statussen en statusovergangen werden gebruikt die niet in dit procesmodel vermeld waren. Daarnaast zagen we in onze dataset dat bepaalde statussen in de praktijk heel weinig gebruikt werden. De wens vanuit de business was om bij de geplande inrichting van een nieuwe versie van het facilitaire systeem eveneens het proces te herontwerpen.

Concluderend zien we dat process mining voordelen biedt op meerdere vlakken: Het is gebaseerd op feitelijke informatie (evidence-based) en daarmee kunnen we onderbouwde conclusies trekken over hoe onze processen in werkelijkheid verlopen. Hiermee tonen we verspilling in het proces aan, en zien we waar het proces afwijkt van de gewenste situatie (zoals gedocumenteerd door de business). We zien dat de huidige beschikbare tools de toepassing van process mining laagdrempelig en eenvoudig maken. De uitdaging is om een goede business case voor het uitvoeren van een process mining-project te vinden en de juiste stakeholders uit de business te betrekken. Ten eerste om de issues die door middel van process mining worden gevonden te verklaren, en ten tweede om de voorgestelde verbeteringen te implementeren door het aanpassen van bijvoorbeeld softwareconfiguratie of werkwijzen. We kijken er naar uit om in de toekomst meer systemen onder de loep te nemen en process mining een structureel onderdeel te maken van ons kwaliteitsmanagement.

Remco Bunder

Remco Bunder studeerde Informatie en communicatie technologie en is nu informatie analist bij NS- Stations.

In zijn zoektocht naar innovatie op het gebied van requirements engineering kwam process mining op zijn pad. Vervolgens heeft hij hier een impuls aan gegeven bij de NS. Samen met zijn collega Jacco Vogelsang behaalde hij de 2e plaats tijdens de NS –ICT innovatie awards 2017 met een process mining project. Sindsdien is hij regelmatig betrokken bij proces optimalisatie trajecten met de inzet process mining.

Remco, was ook spreker op het process mining camp 2017 bij de TU-Eindhoven

