

‘TomTom’ voor bedrijfsprocessen

Toepassing van moderne process-miningtechnieken

Aan de hand van een vergelijking tussen bedrijfsinformatie- en navigatiesystemen wordt geïllustreerd hoe informatiesystemen nog onvoldoende inzicht bieden in de lopende bedrijfsprocessen. Moderne process-mininggereedschappen zoals ProM kunnen uitkomst bieden.

Wil van der Aalst

Ondanks het grote belang van efficiënte en effectieve bedrijfsprocessen leveren informatiesystemen onvoldoende inzicht in de lopende processen. Om dit te illustreren wordt de functionaliteit van een typisch bedrijfsinformatiesysteem vergeleken met die van een typisch navigatiesysteem. Deze vergelijking laat zien dat in de meeste organisaties een fatsoenlijke landkaart ontbreekt, dat file-informatie afwezig is, dat de aankomsttijd onbekend is en dat afwijkingen van de route niet toegestaan zijn. Daarom een pleidooi voor ‘TomTom-functionaliteit’ voor bedrijfsprocessen. Veel van de gewenste functionaliteit kan worden gerealiseerd met moderne process-miningtechnieken.

Procesnavigatie

Dankzij het enorme succes van TomTom beschikken veel Nederlandse gezinnen over een navigatiesysteem in de auto. Het is verbazingwekkend hoe snel mensen gewend zijn geraakt aan deze functionaliteit en er niet meer zonder kunnen. Omdat vaak vergelijkingen worden gemaakt tussen verkeersstromen op de weg en werkstromen in grote organisaties, is het interessant de functionaliteit van een typisch bedrijfsinformatiesysteem te vergelijken met die van een navigatiesysteem. Hierbij valt een aantal zaken op:

1. *In veel gevallen ontbreekt een goede kaart.* Ondanks het feit dat organisaties veel tijd investeren in het documenteren van de bedrijfsprocessen, zijn de resulterende procesmodellen vaak van bedroevende kwaliteit. De processchema's hebben weinig te maken met de echte processen. Men spreekt daarom ook wel over de ‘PowerPoint-werkelijkheid’.
2. *Het is onmogelijk traploos in en uit te zoomen.* Procesmodellen zijn vaak statische schema's vergelijkbaar met een klassieke landkaart. Hierdoor heeft het model vaak te veel of te weinig detail. In een navigatiesysteem kan men wel in- en uitzoomen en kan de visualisatie worden aangepast op basis van voorkeuren en doelen. Zo kunnen bijvoorbeeld garages, tankstations en hotels op verzoek worden weergegeven.
3. *Als er wel een elektronische kaart bestaat, is deze vaak onnodig dwingend.* Workflow-managementsystemen worden gedreven door gedetailleerde en accurate processchema's. Het probleem is echter dat deze vaak een specifieke procesgang afdwingen. Vergelijk dit

met een navigatiesysteem dat de uiteindelijke keuze niet aan de gebruiker laat en de bestuurder bijvoorbeeld dwingt een afgesloten straat in te rijden. Soms willen medewerkers om goede redenen afwijken en moet het systeem dit toestaan.

4. *De kaart laat geen file-informatie zien.* Procesmodellen zijn vaak papieren tijgers. Ze komen alleen tot leven als er actuele informatie op geprojecteerd wordt. Zo is het buitengewoon nuttig een kaart te zien met file-informatie. Waar in een traditionele fabriek de knelpunten duidelijk zichtbaar zijn, worden in moderne bedrijfsinformatiesystemen deze 'files' echter vaak onzichtbaar gemaakt.
5. *De route kan niet worden herberekend.* Ook als er duidelijke verstoringen optreden (bijvoorbeeld lange files) en gebruikers proberen om te rijden, blijven informatiesystemen meestal stoïcijns het oorspronkelijke procesmodel volgen. Bedrijfsprocessen zijn soms uitermate dynamisch en het zou beter zijn als de informatiesystemen hierop kunnen inspelen.
6. *De aankomsttijd en locatie zijn onbekend.* Op elk moment geeft het navigatiesysteem aan wat de nog af te leggen afstand is en wat de verwachte aankomsttijd is. Ook deze functionaliteit ontbreekt in de huidige generatie bedrijfsinformatiesystemen. Vaak is er weinig inzicht in hoelang zaken al in het systeem zitten en wanneer ze verwacht worden klaar te zijn. Soms is zelfs onduidelijk waar in het proces een concreet geval zit; er is geen gps-informatie beschikbaar.

Informatiesystemen zijn dus erg beperkt als het gaat om 'TomTom functionaliteit'. Deze analyse is dan ook bedoeld om de lezer te prikkelen en te laten nadenken over navigatiefunctie in bedrijfsprocessen.

En business intelligence dan?

Business intelligence (BI) is een kreet die door leveranciers wordt gebruikt om te verwijzen naar software die kan worden gebruikt om gegevens over operationele bedrijfsprocessen te verzamelen en deze vervolgens te analyseren. Het doel van BI-software is om meer kennis en inzicht te verkrijgen die kan worden gebruikt om de processen beter te besturen en in te richten. De focus ligt hierbij vaak op performance-indicatoren die betrekking hebben op kosten, doorlooptijden, serviceniveaus en bezettingsgraden. De 'I' in BI suggereert dat deze systemen intelligent zijn. Niets is minder waar. Vaak gaat het om vrij triviale analyses van gestructureerde data. Veel processen zijn lastig te karakteriseren in termen van performance-indicatoren omdat de werkelijkheid niet te vangen is in een paar getallen. Ook veronderstellen BI-tools gestructureerde processen en gestructureerde gegevens waarin geen afwijkingen mogelijk zijn.

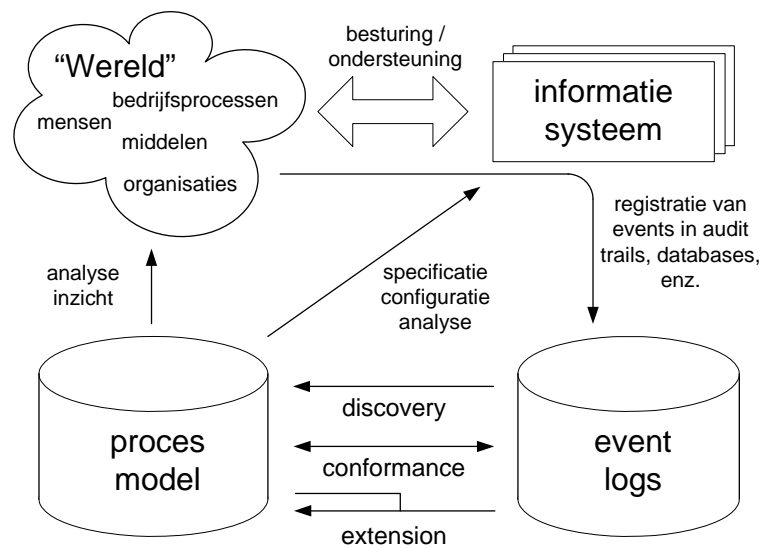
Neem bijvoorbeeld de behandeling van hartpatiënten in een ziekenhuis. Met behulp van de huidige generatie BI-tools is het mogelijk om vast te stellen wat het aantal operaties is geweest in een bepaalde periode en hoelang de gemiddelde behandeling van de patiënten na de operatie heeft geduurd. Ook is het mogelijk om één individuele patiënt te volgen. Helaas is het niet mogelijk om af te leiden hoe de processen daadwerkelijk verlopen zijn en waar en waarom er afwijkingen en knelpunten ontstaan. De reden hiervoor is dat de werkelijkheid wordt gevangen in getallen zonder dat het echte onderliggende proces adequaat zichtbaar gemaakt wordt. De 'managementdashboards' laten meters en grafieken zien, maar niet het onderliggende proces. De meer geavanceerde BI-tools beelden performance-indicatoren af op een procesmodel. Het gaat dan echter altijd om een vooraf gedefinieerd procesmodel dat niet noodzakelijk de werkelijkheid hoeft weer te geven. Vaak laten met de hand gemaakte procesmodellen een sterk geïdealiseerde wereld zien waardoor de daadwerkelijke verbeterpunten verborgen blijven.

Kortom, BI-tools zijn niet intelligent en bieden geen ‘TomTom-functionaliteit’. Gelukkig zijn er recent op het gebied van *process mining* belangrijke doorbraken gerealiseerd die kunnen helpen bij het realiseren van innovatieve navigatiefuncties.

Process mining

Om het concept van process mining duidelijk te maken is het van belang te realiseren dat informatiesystemen vaak een afspiegeling van de werkelijkheid vormen. Enerzijds worden informatiesystemen gedreven door modellen van de bedrijfsprocessen die ze ondersteunen. Anderzijds is er een voortdurende interactie tussen het bedrijfsproces en het informatiesysteem. Neem bijvoorbeeld een ERP-systeem als SAP R/3. Dit systeem veronderstelt dat binnen organisaties bepaalde processen plaatsvinden, bijvoorbeeld het plaatsen van bestellingen, en dat voor deze processen bepaalde activiteiten in een zinvolle volgorde worden uitgevoerd. Dit zijn de processen die vastgelegd zijn in het zogenaamde SAP-referentiemodel. De ontwikkeling en configuratie van SAP R/3 wordt dus gedreven door deze veronderstelde processen. In werkelijkheid laat het systeem echter ook andere processen en procesvolgordes toe om de benodigde flexibiliteit te realiseren. Onafhankelijk van de wijze waarop het systeem wordt gebruikt, worden zogenaamde *events* gelogd. Een event is een gebeurtenis die overeenkomt met een bepaalde actie of handeling. In de context van informatiesystemen wordt ook wel gesproken over (database)*transactions* of *audit trail entries*. De *event logs* van systemen als SAP R/3 laten een afbeelding van de werkelijkheid zien en vormen het startpunt voor process mining.

Het is van belang om in te zien dat in toenemende mate gebeurtenissen in bedrijfsprocessen geregistreerd worden in event logs. Pakketten voor het realiseren van informatiesystemen, bijvoorbeeld enterprise resource planning (ERP)-, customer relationship management (CRM)-, product data management (PDM)- en workflow management (WfM)-systemen, bieden standaard de functionaliteit om gedetailleerde event logs te genereren. Ook hebben steeds meer systemen ingebouwde loggingfunctionaliteit. Complexe professionele systemen zoals medische apparatuur (bijvoorbeeld de röntgensystemen van Philips), high-end kopieermachines, lithografiesystemen (bijvoorbeeld de wafer steppers van ASML) zijn in staat alle handelingen te registreren en te distribueren via internet. Ook het toenemende gebruik van webservices leidt tot steeds meer en betere event logs. Door bijvoorbeeld SOAP-messages te volgen kan de communicatie tussen partijen in een SOA (service-oriented architecture) in kaart worden gebracht.



Figuur 1. Drie soorten process mining: discovery, conformance en extension

Drie soorten process mining

De brede beschikbaarheid van event logs en de koppeling tussen bedrijfsprocessen en informatiesystemen maakt process mining mogelijk. Zoals aangegeven in figuur 1 beschouwen we drie soorten process mining: *discovery*, *conformance* en *extension*.

Process-miningtechnieken gericht op *discovery* leiden automatisch modellen af uit de event logs. Uit ruwe informatie kunnen bijvoorbeeld procesmodellen worden gedestilleerd die weergeven wat er daadwerkelijk heeft plaatsgevonden. Het is dus niet de analist die met hand een (veelal subjectief) model van de werkelijkheid maakt. In plaats daarvan wordt het event log onderzocht op mogelijke causale verbanden, waarna de gevonden verbanden worden samengevat in een grafisch model (bijvoorbeeld een Petrinet-achtige notatie zoals gebruikt in tools als Protos, COSA en YAWL of de EPC-notatie gebruikt in tools als ARIS en SAP R/3). Dit klinkt utopisch. Recent zijn er echter krachtige technieken ontwikkeld die dit mogelijk maken. De discoverytechnieken beperken zich niet tot het procesaspect (dat wil zeggen de volgorde van activiteiten). Indien events ook informatie bevatten over resources (wie heeft de activiteit uitgevoerd en welke middelen zijn hierbij gebruikt), data (wat waren de karakteristieke kenmerken), tijd (wanneer vond de activiteit plaats), kunnen ook andere aspecten afgedekt worden. Het is bijvoorbeeld mogelijk om op basis van event logs een zogenaamd sociaal netwerk af te leiden. In een sociaal netwerk worden de personen en/of organisatieonderdelen weergegeven met hun onderlinge relaties. Hierdoor is het mogelijk weer te geven dat sommige personen een centrale plaats innemen in de organisatie en dat er veel werk van de ene naar de andere afdeling stroomt.

Een tweede categorie van process-miningtechnieken is gericht op *conformance*. Het doel is objectief te toetsen of de werkelijkheid overeenkomt met de verwachting of vooraf opgestelde eisen. *Conformance checking* kan helpen bij het concretiseren van *compliance* en nieuwe regelgeving zoals Sarbanes-Oxley (SOx). Schandalen binnen organisaties als WorldCom, Enron, AOL, Tyco en Qwest laten zien dat de werkelijke procesgang sterk kan afwijken van veronderstelde processen. Binnen veel organisaties bestaan er procesbeschrijvingen die weergeven hoe processen uitgevoerd (dienen te) worden. Deze procesmodellen worden vaak gemaakt met eenvoudige tekentools zoals Visio of met meer geavanceerde gereedschappen

zoals Protos en ARIS. Er is echter geen enkele garantie dat de processen ook daadwerkelijk deze modellen volgen. In veel gevallen beschrijven procesmodellen een geïdealiseerde 'PowerPoint-werkelijkheid' die ver afstaat van de dagelijkse realiteit. Daarom is het zinvol om op basis van event logs deze modellen te toetsen. Process-miningtechnieken kunnen aangeven hoe goed de 'fit' tussen realiteit en model is en kunnen ook weergeven waar in het proces de belangrijkste afwijkingen plaatsvinden. Vervolgens kunnen de gevallen die afwijken verder worden onderzocht met de eerdergenoemde discoverytechnieken. Ook is het mogelijk zogenaamde *business rules* te toetsen. Het is bijvoorbeeld interessant te weten in hoeveel procent van de gevallen een business rule daadwerkelijk gevolgd wordt.

De laatste categorie, *extension*, pakt een al bestaand model en vult dit aan met gegevens uit het event log. Het initiële model kan met de hand gemaakt zijn, maar kan ook ontdekt zijn door middel van process mining. Een voorbeeld is het projecteren van knelpunten op een procesmodel op basis van in het log geregistreerde tijdstempels. Ook kan op basis van gegevens in het log bepaald worden wanneer er een bepaald pad in het procesmodel gevolgd wordt. Door middel van klassieke dataminingtechnieken kan bijvoorbeeld ontdekt worden dat een bepaalde categorie klanten vaak een bepaalde tak van het proces doorloopt. Het ultieme doel van process mining is een model te creëren dat goed overeenkomt met de werkelijkheid en dat alle aspecten afdekt (activiteiten, resources, tijd, data enzovoorts).

De weerbarstige werkelijkheid

Aan de Technische Universiteit Eindhoven wordt sinds 2000 onderzoek gedaan naar process mining. Daarvoor lag de nadruk sterk op het modelleren van processen voor workflowsystemen en de analyse van bedrijfsprocessen door middel van simulatie. Door de ervaringen op het gebied van simulatie en het realiseren van workflowsystemen werd echter steeds duidelijker dat er meestal een grote discrepantie is tussen model en werkelijkheid. Helaas hebben de meeste consultants en leveranciers vooral aandacht voor (her)ontwerp en wordt er te weinig energie besteed aan een gedegen analyse van de werkelijkheid waarin uitzonderingen vaak de regel zijn. Dit wordt bijvoorbeeld duidelijk in simulatiestudies waarin vaak aan parameterwaarden moet worden 'gesleuteld' om doorlooptijden te laten overeenkomen met de werkelijkheid.

Ook kan het falen van veel workflowprojecten worden verklaard door naïeve procesmodellen. Ervaringen met de analyse van patiëntenstromen in ziekenhuizen laten bijvoorbeeld zien dat zelfs voor homogene groepen patiënten de processen minder structuur hebben dan verwacht. Bovendien hebben de betrokken partijen vaak uiteenlopende beelden van het proces. Managers zijn vaak niet op de hoogte van de feitelijke procesgang ('helikopterview') en uitvoerenden zien vaak maar een beperkt deel van het proces ('tunnelvisie'). Het sterk vertekende beeld leidt tot systemen die onbruikbaar zijn en tot gemiste kansen om de processen te verbeteren. Zoals eerder aangegeven schieten BI-tools hier tekort doordat ze ook uitgaan van een geïdealiseerde en sterk versimpelde werkelijkheid. Performance-indicatoren op een sterk geaggregeerd niveau zullen ziekenhuizen bijvoorbeeld niet helpen bij het in kaart brengen van en grip krijgen op patiëntenstromen.

BI-tools schieten tekort doordat ze uitgaan van een geïdealiseerde en sterk versimpelde werkelijkheid

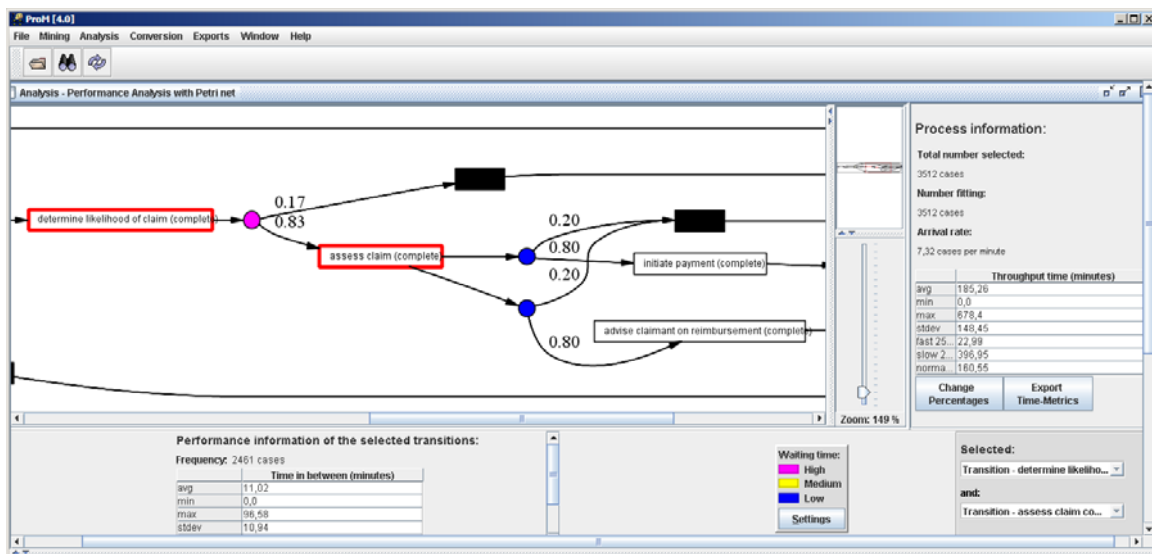
Processen ontdekken met ProM

Gedreven door de geschetste problemen werkt een vijftiental onderzoekers aan de Technische Universiteit Eindhoven aan de ontwikkeling van ProM. ProM is een raamwerk waarbinnen gereedschappen zijn ontwikkeld die de eerdergenoemde vormen van process mining ondersteunen. Er zijn bijvoorbeeld meer dan twintig verschillende discoverytechnieken ontwikkeld.

Het gebruik van ProM start met een zogenaamd MXML-log. Dit is een log waarin informatie over events op een gestructureerde manier wordt opgeslagen. Omdat de meeste informatiesystemen hun informatie niet opslaan in MXML-formaat maar in leveranciersspecifieke databases, audit trails en transactielogs, is het gereedschap ProMimport ontwikkeld. Hiermee kunnen logs vanuit systemen als FLOWer, Staffware, Apache, CPN Tools, CVS, Eastman, PeopleSoft, Subversion, WebSphere en Adept geladen worden. ProMimport maakt het eenvoudig om de benodigde importfunctionaliteit te ontwikkelen en ondersteunt ook het anonimiseren van vertrouwelijke gegevens (bijvoorbeeld namen van medewerkers). Met ProM is het ook mogelijk vele verschillende modeltypen door elkaar te gebruiken. De toolset ondersteunt bijvoorbeeld niet alleen diverse Petrinet-formaten, maar ook EPC-modellen en BPEL-specificaties. ProM kan ook modellen uitwisselen met systemen als Protos, ARIS, Websphere, YAWL en EPC Tools. Daarnaast zijn modelconversies mogelijk. Als met behulp van process mining bijvoorbeeld een Petrinet ontdekt wordt, kan dit model vertaald worden naar BPEL, EPC's en YAWL.

In figuur 2 is een schermafbeelding te zien van ProM waarin op basis van een log een procesmodel is afgeleid. Het procesmodel is afgebeeld als een EPC die kan worden ingeladen in ARIS, een Petrinet dat als invoer kan dienen van een scala van analysegereedschappen en een YAWL-model dat meteen kan worden uitgevoerd door het open-source workflowsysteem YAWL.

ProM ondersteunt niet alleen discovery maar ook de twee andere vormen van process mining: conformance en extension. De conformance checker kan bijvoorbeeld in een procesmodel grafisch laten zien waar de grootste afwijkingen plaatsvinden en deze afwijkingen ook kwantificeren. Figuur 3 laat een voorbeeld zien van extension. Hier is in een procesmodel aangegeven wat de doorlooptijd is, waar de knelpunten zitten en hoeveel tijd er gemiddeld tussen twee activiteiten verstrijkt. Deze informatie kan weer worden gebruikt om volledig automatisch simulatiemodellen te genereren.



Figuur 3. ProM kan ook de bottlenecks zichtbaar maken in afgeleide modellen

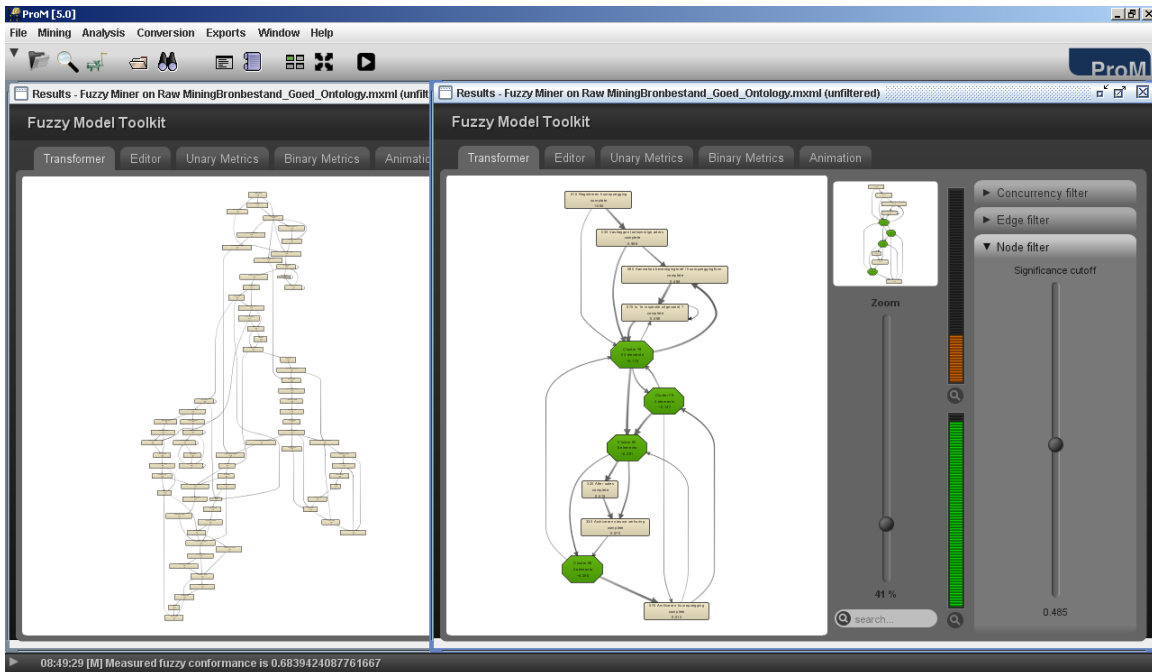
‘TomTom-functionaliteit’ met ProM

Uit de vergelijking tussen navigatiesystemen en bedrijfsinformatiesystemen kwam een lijst met aantal opvallende zaken naar voren. Ter afsluiting kijken we hoe ProM een oplossing biedt voor deze tekortkomingen.

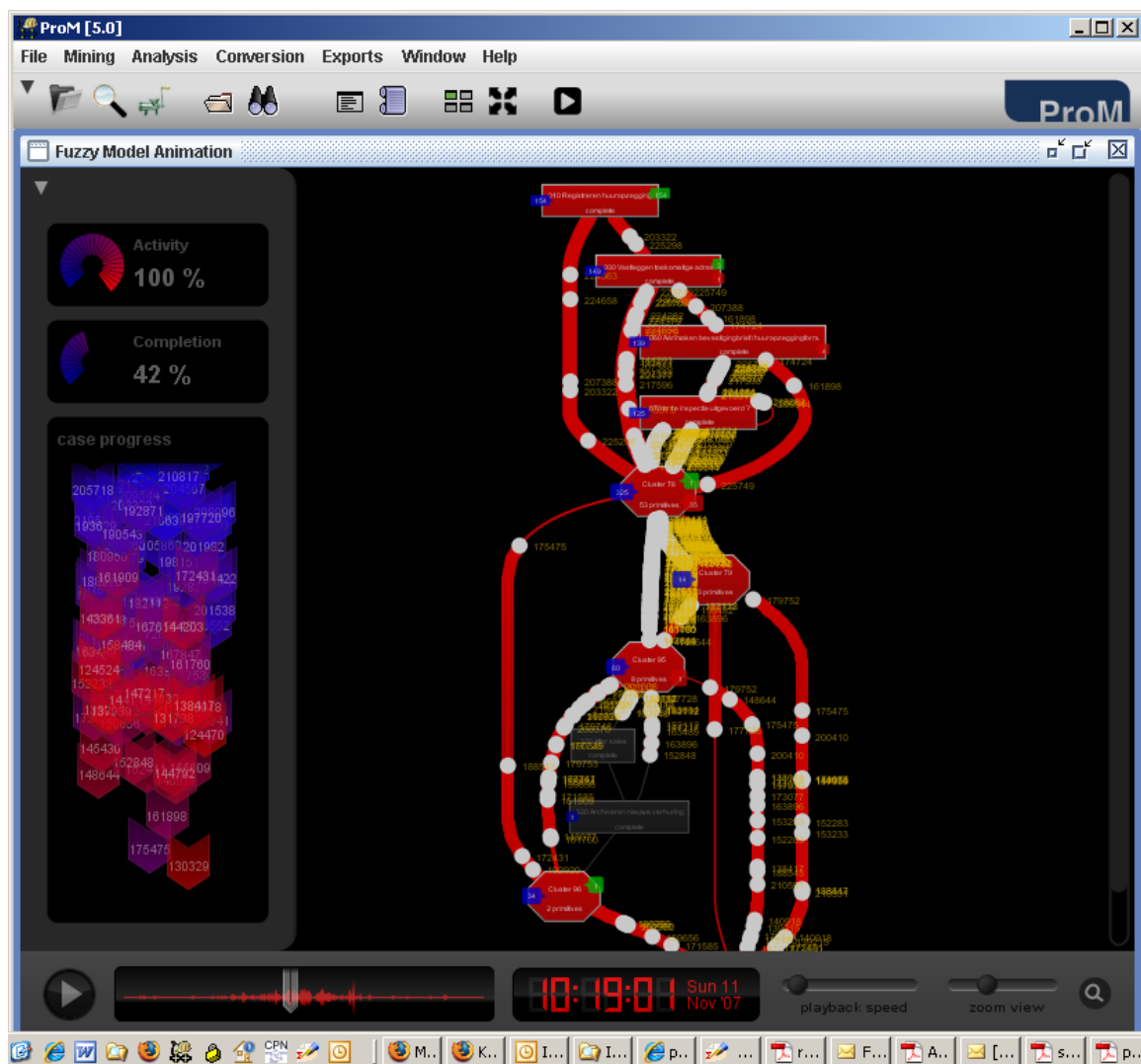
Om te beginnen is het met behulp van process mining zeer wel mogelijk om een goede en accurate kaart van een bedrijfsproces te maken. Door de kaart te baseren op event logs kan de daadwerkelijke procesgang blootgelegd worden. Ook is het mogelijk traploos de detaillering aan te passen door gebruik te maken van technieken in de Fuzzy Miner van ProM. Hier worden op basis van frequenties en correlaties bepaalde paden en activiteiten weggelaten. Net zoals kleinere wegen en dorpen bij uitzoomen door TomTom worden weggelaten of samengenomen, kunnen activiteiten en hun verbindingen worden weggelaten of samengenomen. In figuur 4 is te zien hoe hetzelfde proces met meer of minder detail kan worden weergegeven.

Figuur 5 laat zien dan het mogelijk is om historische en actuele informatie te projecteren op procesmodellen. Dit is vergelijkbaar met het visualiseren van file-informatie, maar kan ook nog veel verder gaan en knelpunten en andere vaak voorkomende problemen visualiseren. Workflow-managementsystemen, ERP-systemen en andere systemen voor procesondersteuning hebben de eigenschap bepaalde procedures dwangmatig op te leggen. De combinatie van flexibele workflowtechnologie en process mining kan worden gebruikt voor een vorm van procesondersteuning die niet noodzakelijk dwingend is. Voorbeelden zijn de Recommendation Service en de Prediction Service in ProM. De Recommendation Service geeft advies op basis van historische informatie. In de werkbak (ook wel *inbox* genoemd) van de gebruiker worden meerdere alternatieven aangeboden. Deze zijn echter geordend op basis van criteria als verwachte kosten of doorlooptijd. De Prediction Service gebruikt een soortgelijk mechanisme. Nu ligt de nadruk echter uitsluitend op het voorspellen van een bepaalde eigenschap zoals resterende doorlooptijd of succeskans. Dit is vergelijkbaar met de verwachte aankomsttijd weergegeven in een navigatiesysteem. Verder is het zo dat op basis van nieuwe informatie de Recommendation Service en de Prediction Service de situatie steeds weer opnieuw beoordelen, ook nadat de gebruiker bewust is afgeweken van de normale route.

Op deze wijze kan beter worden omgesprongen met uitzonderingen, sterk dynamische processen en veranderende omstandigheden.



Figuur 4. Hetzelfde proces met veel detail (links) en weinig detail (rechts)



Figuur 5. Procesmodel met file-informatie

ProM

ProM is ontwikkeld onder een open-source licentie (CPL) en wordt gedistribueerd via www.processmining.org. De ontwikkeling van ProM wordt geleid vanuit de Technische Universiteit Eindhoven, maar er wordt gebruik gemaakt van een breed internationaal netwerk van samenwerkende onderzoeksgroepen op het gebied van process mining. Naast het aanbieden van een zeer toepasbaar gereedschap is het doel van ProM ook om leveranciers van BI- en BPM-tools te stimuleren meer ‘werkelijk intelligente’ functionaliteit te bieden. BPM'one van Pallas Athena, Futura Reflect van Futura Technology, ARIS van IDS Scheer en de Discovery & Visualization Service van Fujitsu zijn schaarse voorbeelden van commerciële gereedschappen die op dit moment enige process-miningfunctionaliteit aanbieden. Zie www.processmining.org voor meer informatie.

Prof. dr. ir. Wil van der Aalst

is hoogleraar Informatiesystemen aan de faculteit Wiskunde & Informatica van de Technische Universiteit Eindhoven en is als deeltijdhoogleraar verbonden aan de faculteit Technologie Management van dezelfde universiteit en de Business Process Management-groep van Queensland University of Technology in Australië. E-mail: w.m.p.v.d.aalst@tue.nl. WWW: www.vdaalst.com.