

# Behoeftte aan visuele technieken voor verbetering controle

Onderzoek naar de mogelijkheden om visualisaties en heuristieken in te zetten voor analyse van financiële gegevens

Quintra Rijnders, Thom Eijken, Marcia Fissette en Jori van Schijndel

**SAMENVATTING** Het controleren van financiële overzichten is één van de kerntaken van accountants. Deze overzichten zijn veelal tot stand gekomen op basis van grote hoeveelheden journaalposten. De huidige stand van de techniek maakt het in theorie mogelijk accountants te equiperen met visuele middelen op basis waarvan zij inzicht kunnen krijgen in het grootboek als geheel. Voor managementinformatie binnen bedrijven signaleren wij een toenemend gebruik van visuele middelen. Toch staat het gebruik van deze visuele middelen binnen de accountantscontrole nog in de kinderschoenen. Dit artikel beschrijft in dat kader een praktijkonderzoek naar de inzet van een visuele methode gedurende de accountantscontrole.

**RELEVANTIE VOOR DE PRAKTIJK** De behoefte aan innovatie bij openbaar accountants is groter dan ooit. Er zijn maatschappelijke discussies over de toegevoegde waarde van de controlerend accountant. De traditionele controleaanpak staat ter discussie, terwijl op technisch gebied steeds meer mogelijk is. Op het gebied van managementinformatie zien wij een trend in het toenemende gebruik van grafisch ingerichte informatie ten koste van 'ouderwetse' spreadsheets. Grafische informatie blijkt onze eigen kloksnelheid enorm te verhogen, neem de tweejarige die zonder moeite een tablet bedient. Accountants werken vaak nog met grootboekkaarten, consolidatiestatens en prints met periode 13-boekingens. Hiermee komt de vraag op wat visuele technieken voor accountants zouden kunnen betekenen.

## 1 Inleiding

Grootboekadministraties bestaan in onze ervaring tegenwoordig uit (honderd)duizenden tot miljoenen mutaties<sup>1</sup> met een omvang van meerdere gigabytes. Robert Birge schatte in 1996 dat een mens waarschijnlijk een geheugencapaciteit van drieduizend gigabytes heeft.<sup>2</sup> Tot zover geen probleem. Onderzoek door Thomas Landauer (1986) heeft echter uitgewezen dat

ons brein ongeveer twee bytes informatie per seconde opneemt. Het op mutatiebasis doorgronden van een grootboek binnen enkele dagen of zelfs weken lijkt dus een onmogelijke opgave. Na vier weken zonder eten en slapen zou de teller pas op 0,005 gigabyte staan.

Een stukje geschiedenis. Een anekdote is dat dubbel boekhouden een van de verworvenheden is die Marco Polo<sup>3</sup> van zijn verre reizen meenam. Toen Luca Pacioli in 1494 een boek erover schreef, kreeg dubbel boekhouden bekendheid. Het doel van (dubbel) boekhouden is om op elk moment inzicht te kunnen verkrijgen in de financiële positie, vermogen en resultaat (Van Zanten, 1892). Sinds 1494 is de bedrijfsvoering behoorlijk veranderd als gevolg van globalisatie van markten en doordat bedrijfsprocessen zijn verveelvoudigd, complex zijn geworden en zijn gedigitaliseerd. Dit resulteert in de vele mutaties uit de eerste alinea.

De maatschappij verwacht dat fraude en fouten door accountants worden ontdekt.<sup>4</sup> Maar kan zij dit redelijkerwijs verwachten? Boekhoud- en fraudeschandalen katalyseren ontwikkeling van regelgeving, die in toenemende mate vraagt om detailcontroles en checklists (onder andere geïllustreerd door het artikel van Ter Hoeven, 2005). Het grootboek groeide uit zijn traditionele doorslagjasje en het magentakleurige potlood van de accountant werd vervangen door een rode ballpoint. Accountants worstelen met het in balans houden van een op risico's gebaseerde jaarrekeningcontrole, budgetdruk en toereikende detailtests (Ter Hoeven, 2005). Is de verwachtingskloof (Wallage, 2005; Has-sink, 2002) niet te groot geworden?

In maart 2011 hebben wij (Rijnders et al., 2011) onderzoek gedaan naar het toepassen van Bayesiaanse netwerktechnieken<sup>5</sup> op een grootboekadministratie. In dit artikel stelden wij aan de hand van onder andere werk

van Schouten (2007) en De Haas (2003) dat het in kaart brengen van opvallende transacties in het kader van de accountantscontrole vaak gedaan wordt vanuit een traditionele benadering van een kwalitatieve risicoanalyse en een kwantitatieve (data-)analyse gebaseerd op ervaringsregels. Daarnaast staat het gebruik van visuele technieken binnen de accountantscontrole in de kinderschoenen (Mueller-Wickop et al., 2011). Risico-indicatoren die niet betrokken zijn in voornoemde kwalitatieve risico- en data-analyse worden buiten beschouwing gelaten. Dit is risicovol in zichzelf, omdat hierbij wordt uitgegaan van het bestaande referentiekader van de controlerend accountant. Terwijl waardevolle informatie, aanwezig in de onderliggende data van de te controleren financiële overzichten buiten beschouwing wordt gelaten.

Met technieken uit de kunstmatige intelligentie, hebben wij onderzocht of data in het grootboek op radicaal andere wijze geanalyseerd kon worden om tot nieuwe inzichten te leiden. Aan de hand van voorgangers op dit gebied (Koopmans, 2005; Veenstra & Heertje, 2005) kozen wij voor het gebruik van Bayesiaanse netwerken. Deze Bayesiaanse netwerken analyseerden de grootboekadministratie regel voor regel. Dit leidde niet tot direct toepasbare resultaten, waardoor vervolgonderzoek nodig was. Als suggestie gaven wij het doen van vervolgonderzoek naar de vraag of software volgtijdelijke relaties in grootboekregels zou kunnen vinden. In dit artikel gaan wij in op het onderzoek op dit gebied dat wij sinds 2011 hebben verricht. Onze hypothese bij dit onderzoek is dat het inzetten van visuele technieken in de accountantscontrole waarde toevoegt aan de kwaliteit en efficiëntie van de jaarrekeningcontrole. Met visuele technieken bedoelen wij technieken die de accountant in staat stellen om in één voor het menselijk oog behapbaar overzicht de stromen in het grootboek te kunnen zien. Aan de hand van grootboekvisualisaties onderzoeken wij met controlerend accountants of en op welke manier zij meerwaarde op kwaliteit en efficiëntie ervaren van visuele technieken.

In paragraaf 2 beschrijven wij het theoretisch raamwerk voor onze methode en de gebruikte visualisatie-technieken. In paragraaf 3 is de onderzoeksmethode beschreven. De resultaten van ons onderzoek zijn beschreven in paragraaf 4. Onze conclusie en enkele suggesties voor verder onderzoek zijn weergegeven in paragraaf 5.

## 2 Theoretisch raamwerk

Bij ons vorige onderzoek werd voor elke mutatie in het grootboek individueel bepaald of deze mogelijk afwijkend was ten opzichte van de rest van het grootboek. Onze methode is daarmee dus wel in staat om opvallende mutaties te vinden, maar deze afwijkingen be-

perken zich tot karakteristieken van de mutatie, bijvoorbeeld het bedrag of het moment van de mutatie. Echter, in veel gevallen kan op basis van de karakteristieken van één enkele grootboekmutatie niet worden bepaald of deze nader bekeken dient te worden. Zeker bij fraude zal vaak geprobeerd worden om door middel van valse of misleidende registraties te verhullen wat er werkelijk gebeurt en indien goed opgezet, zullen individuele registraties niet buiten de norm vallen en dus niet opvallen (Lari, 2009). Het is vaak juist de combinatie van meerdere mutaties binnen een boekingsgang die bepaalt of deze mutaties interessant zijn voor nader onderzoek.

Dat juist een combinatie van mutaties interessant is, ligt in het feit dat het grootboek de resultante is van financiële bedrijfsprocessen. Ieder proces (bijvoorbeeld inkoop of verkoop) leidt tot financiële stromen in de grootboekstroom: een aantal opeenvolgende grootboekmutaties die betrekking hebben op een reeks grootboekrekeningen. Sommige van deze stromen zijn zeer frequent terwijl andere slechts sporadisch voorkomen. Voor elk type proces hebben accountants vaak een bepaalde, op ervaring gebaseerde norm met betrekking tot de stroom die door het proces tot stand zou moeten komen in het grootboek. Mutaties vallen dan op zodra ze onderdeel zijn van een stroom die afwijkt van deze empirische norm. Het is mogelijk handmatig deze stromen te volgen, maar vanwege de grote hoeveelheid mutaties en de beperkte tijd die meestal beschikbaar is, zullen accountants zich moeten beperken tot een steekproef. Idealiter wil men echter het hele grootboek analyseren om voor elke stroom binnen dit grootboek te bepalen of deze al dan niet conform de verwachting en materieel is en daarom meer aandacht verdient. Er is derhalve behoefte aan een methode om elke stroom binnen het grootboek in kaart te brengen en te kunnen analyseren.

Ten behoeve van het in kaart brengen en analyseren van de door een proces tot stand gekomen gegevens, wordt vaak gebruik gemaakt van een techniek genaamd "process mining" (Van der Aalst, 2011). Deze techniek maakt het mogelijk de processen te identificeren zoals ze werkelijk voorkomen zonder bevooroordeeld te zijn door bijvoorbeeld gebruik te maken van voorgedefinieerde procesmodellen. Process mining is voor zover ons bekend nog niet eerder gebruikt voor de analyse van een grootboek. De waarschijnlijke reden hiervoor is dat om process mining te kunnen toepassen, er eerst een uniek kenmerk moet zijn waarmee mutaties van dezelfde stroom herkend kunnen worden; dit kenmerk is niet in het grootboek aanwezig.

Ondanks globalisering en complexere processen, is het principe van dubbel boekhouden nog steeds hetzelfde.

de: tegenover één of meerdere debetmutaties staan één of meerdere creditmutaties die optellen tot hetzelfde bedrag. Er wordt echter niet geregistreerd tot welk proces een mutatie hoort, welke mutaties bij elkaar horen en welke mutaties elkaar opvolgen. Bij kleine boekhoudingen is dit ook niet nodig, omdat het verloop van de mutaties relatief gemakkelijk te volgen is. Bij moderne grote bedrijfsvoeringen zijn er echter zoveel verschillende grootboekrekeningen, geautomatiseerde bedrijfsprocessen en verschillende mensen betrokken dat een compleet en helder overzicht niet direct paraat is. In dit kader is het ook belangrijk voor een visualisatie om binnen een journaalpost vast te stellen welke mutaties bij elkaar horen. Om een computer in staat te stellen een visualisatie van deze mutaties te maken, moet de computer elke individuele boeking tussen grootboekrekeningen kunnen vaststellen. Om die reden is het vaststellen van stromen op een hoger niveau zoals op dagboekniveau niet toereikend voor het bewerkstelligen van een visualisatie.

Met ons onderzoek exploreren wij de mogelijkheden om processen te extraheren uit het grootboek in de vorm van stromen en deze te visualiseren om zo efficiënt inzicht te verkrijgen in de transactionele data uit het grootboek.

Om te bepalen via welke grootboekrekeningen de stromen lopen zal het hierboven besproken kenmerk moeten worden toegevoegd aan de grootboekdata. Eerst zal vastgesteld moeten worden welke debet- en creditmutaties in een journaalpost bij dezelfde groep behoren. In elke groep van mutaties wordt ten minste één grootboekrekening voor een bepaald bedrag gedebiteerd en ten minste één grootboekrekening gecrediteerd, waarbij de som van de debetbedragen gelijk is aan de som van de creditbedragen. Vaststellen welke debet- en creditmutaties tot dezelfde groep horen kan daarom ook worden beschreven als het koppelen van debetbedragen uit een journaalpost aan creditbedragen in dezelfde journaalpost. Hierin herkennen wij een wiskundig vraagstuk, namelijk het ‘subset sum’-probleem (Martello & Toth, 1985).<sup>6</sup> Om een zo goed mogelijke oplossing te vinden voor dit type probleem maken wij gebruik van heuristieken. Een heuristiek is een intuïtieve vuistregel om een specifiek probleem op te lossen (Michalewicz & Fogel, 2004). Een simpel voorbeeld van een heuristiek is de vuistregel dat een staplengte ongeveer gelijk is aan 1 meter.

Een accountant zal vaak kijken naar een aantal elementen binnen de journaalpost – omschrijving, bedrag, grootboekrekening – om inzicht te krijgen in de aard van de boeking. Accountants gebruiken dus tijdens hun werkzaamheden heuristieken om te bepalen welke mutaties tot dezelfde stroomgroep behoren. Onze

methode, die automatisch deze stroomgroepen bepaalt, gebruikt dergelijke heuristieken.

Wanneer een journaalpost slechts één debet- en één creditmutatie bevat, is het duidelijk dat deze mutaties bij dezelfde stroomgroep horen. Zodra een journaalpost meerdere debet- en creditmutaties bevat weten we niet direct welke van deze mutaties tot dezelfde groep behoren, omdat een journaalpost meerdere groepen van mutaties kan bevatten. Om in deze gevallen te bepalen welke mutaties binnen een journaalpost tot dezelfde stroomgroep behoren, gebruiken wij de heuristiek: hoe dichter de mutaties bij elkaar in een journaalpost voorkomen, hoe waarschijnlijker het is dat zij tot dezelfde stroomgroep behoren. Deze heuristiek is gebaseerd op de aanname dat mutaties in journaalposten die bij elkaar horen in een groep, gelijktijdig worden geregistreerd. Deze aanname geldt niet altijd voor extreem grote journaalposten waarin verschillende mutaties gelijktijdig, bijvoorbeeld via “spreadsheet upload”, geregistreerd zijn (Debreceňy & Gray, 2010). In deze gevallen moet de heuristiek versoepeld worden. De methode koppelt voor deze gevallen niet alleen de debet- en creditmutaties aan elkaar die direct na elkaar geregistreerd zijn, maar koppelt ook mutaties die iets verder van elkaar af staan binnen de journaalpost. Om deze koppeling te maken wordt er gekeken naar diverse karakteristieken van de mutaties (o.a. mutatieomschrijving), waarbij dus gebruik wordt gemaakt van de heuristiek dat mutaties die tot eenzelfde groep behoren, vergelijkbare karakteristieken bevatten.

Door de hiervoor genoemde heuristieken toe te passen, kunnen we binnen een journaalpost de mutaties identificeren die tot dezelfde stroomgroep behoren. Echter, we weten nog niet wat de relatie is tussen mutaties in verschillende journaalposten, waardoor een stroom nog niet van begin tot eind kan worden gevolgd. Bijvoorbeeld, gegeven een of meerdere mutaties die voor een bedrag de grootboekrekening ‘Vorraden’ debiteren, wat zijn de bijbehorende mutaties op de grootboekrekening ‘Vorraden’ die ditzelfde bedrag weer crediteren? Het oplossen van deze vraag kan tevens benaderd worden als het ‘subset sum’-probleem, maar in plaats van journaalposten, zijn nu de debet- en creditmutaties per grootboekrekening relevant. Om die reden zijn ook andere heuristieken nodig die effectief zijn voor het groeperen van mutaties per grootboekrekening: unieke bedragen die slechts één keer voorkomen als debet en credit horen waarschijnlijk bij elkaar. Daarnaast, als een specifiek bedrag meerdere keren wordt gedebiteerd, wordt een creditbedrag gezocht dat overeenkomt met de som van die debetbedragen. Ook bedragen die maandelijks op dezelfde dag worden geboekt, hebben een grotere kans bij elkaar te

horen. Tenslotte, als een kleine groep van mutaties een woord deelt in de tekstuele beschrijving zoals een factuurkenmerk, kenteknummer of een naam, is het waarschijnlijker dat deze bij elkaar horen. Onze methode bepaalt welke van de voornoemde heuristieken in welke volgorde worden toegepast. Na afloop hiervan kunnen de individuele stromen geëxtraheerd worden, door de groepen mutaties per journaalpost te koppelen aan de groepen mutaties per grootboekrekening.

De volgende stap is het inzichtelijk maken van de geïdentificeerde stromen door middel van een visualisatie. Het tekstueel representeren van data heeft een limiet van ongeveer 100 items (Keim, 2002). Naast dat de presentatie van zo'n hoeveelheid tekstuele informatie niet overzichtelijk is, is deze limiet te laag voor het weergeven van de mutaties in een grootboek.

Het doel van de visualisatie is om de relatie die deze mutaties met elkaar hebben, weer te geven. Netwerkdigrammen zijn zeer geschikt voor het weergeven van zowel directe als indirecte relaties (Keller, Eckers & Clarkson, 2006). Visualisatie door middel van netwerkdigrammen maakt het mogelijk om de relaties tussen mutaties in één oogopslag weer te geven. Een belangrijk uitgangspunt hierbij is dat de visualisatie intuïtief moet zijn (Keim et al., 2006; Shneiderman, 1996). Schematisch is derhalve gekozen voor de onderstaande weergave (figuur 1), die laat zien hoe geldmiddelen de organisatie in- en uitstromen.

**Figuur 1** Basismodel voor het visueel weergeven van een grootboek



### 3 Onderzoeksmethode

We hebben de hierboven beschreven concepten geïmplementeerd in een “proof-of-concept”. Dit proof-of-concept werkt volgens een aantal stappen. Eerst maakt een gebruiker een project aan en kan hij de benodigde data inlezen. Daarna wordt de analyse zoals beschreven in paragraaf 2 uitgevoerd. Na afloop daarvan worden de geïdentificeerde stromen teruggekoppeld aan de gebruiker door middel van een visualisatie. De gebruiker kan vervolgens de visualisatie doorzoeken en ervoor kiezen alleen bepaalde details te bekijken door bijvoorbeeld te filteren op materialiteit, aantallen mutaties, gebruikers, grootboekrekening et cetera.

Om te onderzoeken of onze procesanalyse in de praktijk nuttig kan zijn, hebben we zes pilots uitgevoerd op

grootboekdata en achttien gebruikersevaluaties gehouden. Het merendeel van de pilots vond plaats parallel aan de jaarrekeningcontrole (vijf organisaties). Bij deze pilots is de controlerend accountant in staat de gevonden stromen kritisch te analyseren. Door zijn uitgebreide kennis van de gecontroleerde entiteit is de accountant in staat om aan te geven in hoeverre het resultaat van de analyse overeenkomt met zijn beeld van het bedrijf. Daarnaast kan de accountant op basis van de visualisatie afwijkende stromen identificeren en bij het bekijken van de details aangeven of deze daadwerkelijk interessant zijn voor nadere opvolging. Tot slot kan de accountant aangeven of deze werkwijze waarde toevoegt aan de kwaliteit en efficiëntie van de jaarrekeningcontrole.

Voor elke pilot hadden we toegang tot het gehele grootboek voor een enkel financieel boekjaar. Deze data waren voor elke organisatie in eenzelfde digitaal formaat beschikbaar. Om de integriteit van de verwerkte gegevens vast te stellen, is de aansluiting gemaakt met de saldbalans voor het gegeven boekjaar.

De methode werkt met bepaalde informatievelen van de data uit het grootboek. Het gaat hierbij om de informatievelen die interessant zijn voor analyse. Het nummer van de journaalpost is bijvoorbeeld heel nuttig om een journaalpost te identificeren maar totaal oninteressant voor het identificeren van opvallende mutaties. Voor de analyse zijn de interessante informatievelen onder andere het bedrag, grootboekrekening, dagboek, type mutatie, beschrijving, et cetera. Deze subselectie van informatievelen vormt de input voor de analyse.

De “proof-of-concept” is in twee stappen getest:

1. Het bepalen van de stromen door alleen de heuristieken te gebruiken die binnen journaalposten de debets en credits met elkaar verbinden. Deze analyse is getest op één grootboek en geëvalueerd met elf gebruikers. Het doel van de evaluatie in deze stap was om te bepalen of de gebruikers met behulp van een visualisatie inzicht konden krijgen in de grootboekdata. In stap 1 van de evaluatie is de methode geëvalueerd met elf gebruikers, onderverdeeld in twee groepen. De eerste groep bestond uit zes gebruikers met beperkte kennis van financiële gegevens en de tweede groep bestond uit vijf gebruikers met relatief veel kennis<sup>7</sup> van financiële gegevens. De verwachting vooraf was dat de gebruikers met relatief veel financiële kennis, meer inzicht zouden krijgen op basis van de visualisatie.
2. Het toepassen van de volledige methode die gebruik maakt van alle beschikbare heuristieken op vijf grootboeken. Deze stap is uitgevoerd parallel aan de jaarrekeningcontrole van de desbetreffende klanten en is geëvalueerd met de desbetreffende accountants.



Het doel van de evaluatie in deze stap was om te bepalen of de methode waarde toevoegt aan de kwaliteit en efficiëntie van de jaarrekeningcontrole en welke functionaliteiten eventueel nog ontbraken in de methode.

Het grootboek van stap 1 bestaat uit 7.853 journaalposten die samen 59.968 mutaties bevatten. Deze set bevat 34 journaalposten die één creditbedrag of één debetbedrag bevatten. Er zijn 36 journaalposten waarbij de som van de debetbedragen niet gelijk is aan de som van de creditbedragen, deze worden daarom niet meegenomen in de analyse.<sup>8</sup> Twee journaalposten bevatten de openingsbalans. De heuristieken zelf maken geen gebruik van de openingsbalans, maar de methode kan deze wel gebruiken om deze als onderdeel van de visualisatie weer te geven.

De vijf grootboeken die in stap 2 zijn gebruikt bevatten tussen de 20.000 en 3.500.000 mutaties. De resultaten uit bovenstaande onderzoeksmethodiek zullen in de resultatensectie nader toegelicht worden.

Vanwege de huidige fase van het project, is gekozen voor een empirische evaluatie van de resultaten. De reden hiervoor is dat we vooral willen bepalen of de methode bruikbaar is voor de accountants gedurende de jaarrekeningcontrole. Daarnaast wordt steekproefsgewijs onderzocht of de toegepaste heuristieken (beschreven in paragraaf 2) de juiste stromen uit de grootboekmutaties geëxtraheerd hebben.

#### 4 Resultaten

Uit ons onderzoek blijkt dat wanneer alle mutaties die geregistreerd zijn in het grootboek op grootboekrekeningniveau worden gevisualiseerd, stromen onvoldoende inzichtelijk kunnen worden gemaakt. Voor het grootboek dat gebruikt is in stap 1 vindt de methode 1.986 stromen tussen 188 grootboekrekeningen. Dit groot aantal relaties en grootboekrekeningen kan niet in één keer overzichtelijk worden weergegeven. Na indeling van de grootboekrekeningen in groepen, resteren 352 stromen tussen 72 groepen. Hoewel dit aantal tot een betere weergave van de stromen leidt, is het overzicht nog niet voldoende inzichtelijk. Filtermogelijkheden die alleen stromen boven of beneden een geselecteerd bedrag

(bijvoorbeeld materialiteit) of grote aantallen mutaties laten zien, zijn nodig voor beter inzicht.

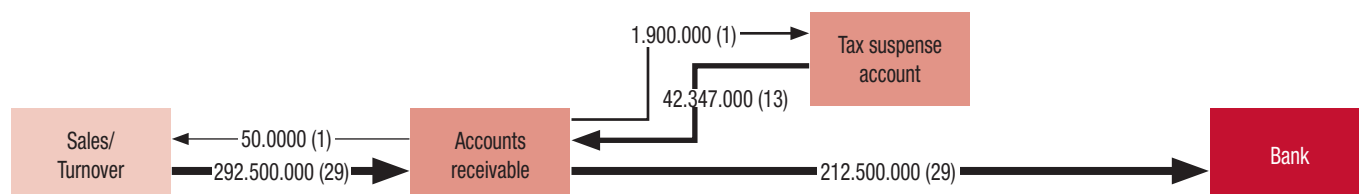
In het overzicht (figuur 2) is de dikte van de stromen tussen de groepen<sup>9</sup> afhankelijk van de hoogte van het bedrag dat van de ene groep van grootboekrekeningen naar de andere is geboekt. Bovendien worden naast deze stromen de bijbehorende bedragen en aantallen individuele boekingen weergegeven wat de inzichtelijkheid bevordert.

In stap 1 van de evaluatie is de methode geëvalueerd met elf gebruikers, onderverdeeld in twee groepen, gebruikers met beperkte en met relatief veel kennis van financiële gegevens. Uit de evaluatie is gebleken dat beide groepen in staat zijn inzicht te krijgen in de grootboekdata door middel van de visualisatie van stromen die bepaald zijn door alleen de heuristieken te gebruiken die binnen journaalposten de debets en credits met elkaar verbinden. Zoals verwacht, verkreeg de tweede groep gebruikers met relatief veel financiële kennis gemiddeld meer inzichten dan de gebruikers in de eerste groep. De testgebruikers in beide groepen van evaluatie hebben aangegeven dat zij de stromen in meer detail willen bekijken.

Deze eerste resultaten geven aan dat het weergeven van de stromen binnen een grootboek op een geaggregeerd niveau inzicht biedt aan gebruikers met zowel weinig als veel kennis van financiële gegevens. Door de vier accountants uit de eerste groep werd opgemerkt dat de visualisatie bruikbaar zou zijn als onderdeel van de risicoanalyse tijdens de planningfase van de jaarrekeningcontrole. Echter, uit de evaluaties blijkt dat de gebruiker behoefte heeft om na het zien van een totaaloverzicht de data verder te verkennen door specifieke stromen meer in detail te bekijken.

Aan de hand van deze resultaten hebben wij de interactie die de gebruiker met de methode heeft, herzien. Belangrijk is dat elke analyse start met een relatief begrijpelijk overzicht. Vanaf dat punt kan de gebruiker interactief details uitzoeken voor juist die stromen die interessant lijken. Om dit te bereiken passen we in stap 2 van de evaluatie extra heuristieken toe voor het herkennen van individuele stromen (zie figuur 3).

**Figuur 2** Eerste pilot visueel weergegeven grootboek (geanonimiseerd)

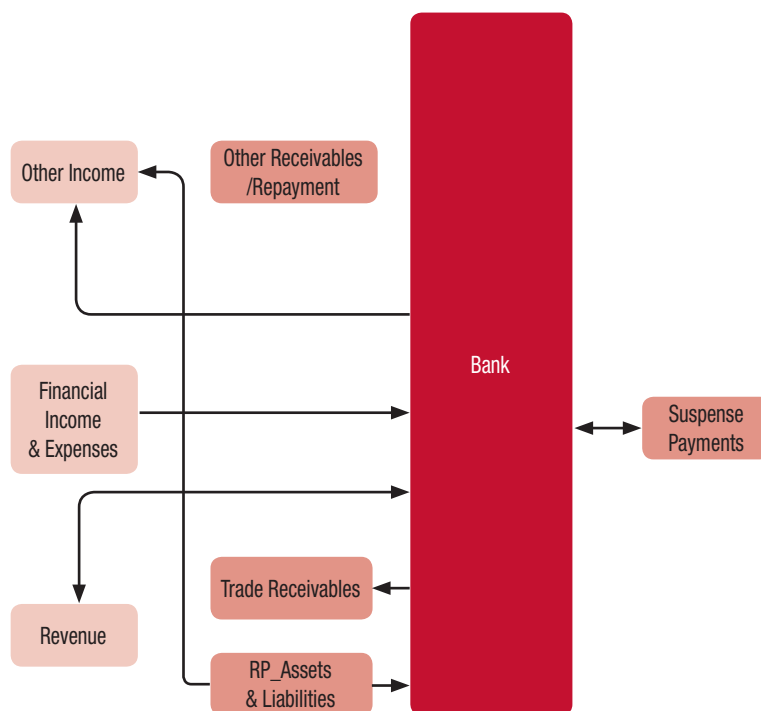


Voor de verschillende datasets in stap 2 zijn er uiteenlopende resultaten. Dit heeft vooral te maken met de verschillen in de manier van boekhouden bij de desbetreffende organisaties. Tekstuele beschrijvingen bij mutaties blijken erg nuttig om mutaties te groeperen en sommige grootboeken hebben ook redelijk informatieve beschrijvingen per regel. Als de beschrijvingen echter ontbreken dan wel te generiek zijn, kan deze informatie niet gebruikt worden.

Het eerste doel van de evaluatie in stap 2 was om te bepalen of de methode waarde toevoegt aan de kwaliteit en efficiëntie van de jaarrekeningcontrole. Tijdens alle evaluaties in stap 2 werden door de accountant vanuit het totaaloverzicht één of meerdere stromen geïdentificeerd die niet overeenkwamen met het verwachtingspatroon van de accountant. Vervolgens zijn de mutaties behorende bij deze stromen in de software met de accountant doorgenomen. In de meeste gevallen bleek dat de desbetreffende mutaties ook tijdens de reguliere werkzaamheden van de controle aan het licht waren gekomen, maar was men met behulp van onze methode in staat om deze op een efficiëntere wijze te vinden dan met 'traditionele' data-analyse. De reden hiervoor is dat het overzicht direct visueel inzicht geeft in de stromen die afwijken van de verwachting en het vinden van de onderliggende details enkele muisklikken kost. Het produceren van dit overzicht is inmiddels een snelle routineklus. Traditionele data-analyse wordt vaak ingezet op delen van het grootboek en kent lang niet altijd een visuele output, waardoor de tijdsvoordelen en inzichten van een algeheel gevisualiseerd grootboek verloren gaan. Dit geeft aan dat door het toepassen van onze methode veel tijd bespaard kan worden, waardoor de investering voor het ontwikkelen van de methode zich vanzelf terugverdient. Daarnaast vond onze methode ook mutaties die tot op dat punt in de jaarrekeningcontrole nog niet ontdekt waren, deze zijn vervolgens door de accountant nader onderzocht.

Tot slot hebben de in de evaluatie betrokken accountants aangegeven dat de verbeterde visualisatie niet alleen bruikbaar is voor de planningsfase van de jaarrekeningcontrole, maar ook voor de balanscontrole, wanneer gegevensgerichte werkzaamheden worden uitgevoerd. Het 'doorklikken' op transactiestromen en uiteindelijk op individuele mutaties stelde de accountants in staat om sneller en gericht detailvragen te stellen aan hun cliënten, subsets van grootboekdata te bewerken om inzicht te verkrijgen en gemakkelijk steekproeven te trekken op posten waar dat doelmatig voor is. Daarnaast hebben de accountants aangegeven dat de visualisatie ook behulpzaam is tijdens de afrondende werkzaamheden, wanneer moet worden vastgesteld of alle standen en stromen binnen het financiële overzicht begrepen en gecontroleerd zijn.

**Figuur 3** Op basis van gebruikersfeedback aangepaste visuele weergave van een grootboek (geanonimiseerd)



Het tweede doel van de evaluatie in stap 2 was om te bepalen welke functionaliteiten eventueel nog ontbraken in de software. Van gebruikers (zowel accountants als functionarissen van de ondernemingen die betrokken zijn geweest in de pilots) kregen we verschillende adviezen teruggekoppeld. De meest belangrijke waren:

- Door verschillende kleuren te gebruiken voor de stromen, kunnen interessante stromen extra naar de voorgrond gebracht worden of kan onderscheid gemaakt worden tussen verschillende typen processen.
- Een bruikbaar filter is om pijlen te verbergen die men verwacht op basis van de typologie van de organisatie, zodat alleen de afwijkende pijlen in het overzicht staan.
- Er is in het algemeen behoefte om meer gebruik te kunnen maken van instellingen die bepalen hoe de visualisatie eruit ziet (bijvoorbeeld door wel/niet weergeven van bedragen, het wel/niet dikker maken van stromen aan de hand van de omvang et cetera).
- Zodra men de details van een stroom bekijkt is er behoefte om de onderliggende mutaties op een overzichtelijkere manier te kunnen zien (bijvoorbeeld door ze te groeperen of samen te vatten, niet alleen als een lijst).

## 5 Conclusie, beschouwing en suggesties voor verder onderzoek

### 5.1 Conclusie en beschouwing

Onze hypothese aan het begin van het onderzoek was

dat het inzetten van visuele methoden in de accountantscontrole waarde toevoegt aan de kwaliteit en efficiëntie van de jaarrekeningcontrole. Uit ons onderzoek blijkt dat visualisatie van data waarde toevoegt voor een accountant. Enerzijds omdat opvallende stromen/mutaties binnen het grootboek sneller en efficiënter worden ontdekt, anderzijds omdat opvallende stromen/mutaties binnen het grootboek worden ontdekt die met 'traditionele' controletechnieken (mogelijk) niet aan het licht komen.

Op basis van dit onderzoek blijkt dat visualisatietechnieken alleen niet toereikend zijn, maar dat accountants ook behoefte hebben aan filter- en andere analysemogelijkheden als de data eenmaal is gevisualiseerd. Visualisatietechnieken moeten het mogelijk maken om in één oogopslag te zien welke mutaties meer aandacht verdienen. Wiskundige wetmatigheden en data-analyse kunnen gedeeltelijk voor controlebewijs zorgen<sup>10</sup> alsmede voor een verfijning van nader te onderzoeken mutaties. Natuurlijk zal het integreren van deze technieken de nodige investeringen met zich meebrengen, maar er is veel inzicht te winnen. De accountant kan gericht op zoek naar fouten en signalen van fraude en behoudt tegelijkertijd overzicht. De doelstelling voor dubbel boekhouden is daarmee gediend: er wordt zowel op geaggregeerd als op detailniveau inzicht verkregen in de financiële positie, het vermogen en het resultaat. Naar onze mening zal de accountantscontrole van de toekomst steeds meer gebruikmaken van visualisatietechnieken, gecombineerd met wiskundige wetmatigheden en data-analyse.

## 5.2 Suggesties voor verder onderzoek

Het verbeteren van visualisatietechnieken in combinatie met het toevoegen van andere data-analyses zal leiden tot bruikbare methodes om snel inzicht te krijgen in de stromen binnen een organisatie.

Een zeer specifieke toekomstige onderzoeksrichting is de toepassing van (semantische) taalanalyse voor het clusteren van mutaties in het grootboek. Momenteel groepeerde de methode debets en credits met eenzelfde waarde, maar bijvoorbeeld correcties op bepaalde boekingen voldoen hier vaak niet aan. De beschrijving bij

een mutatie is dan het enige nuttige aanknooppunt. Een dergelijke taalanalyse zou mogelijk het groeperen van mutaties kunnen faciliteren, als de omschrijving aangeeft dat de mutaties bij elkaar horen, ook al sluiten de bedragen niet direct aan. De huidige vorm van tekstanalyse die we gebruiken is niet verfijnd genoeg om mutatiebeschrijvingen op zo'n inhoudelijke wijze te interpreteren.

Een aanvullende mogelijkheid om de huidige methode te verbeteren is het onderzoek naar integratiemogelijkheden van zelflerende technieken (zoals de Bayesiaanse netwerken die we hebben gebruikt in ons vorige onderzoek) in de huidige visualisatiemethode. Een dergelijke verbeterde methode zou de voordelen van visualisaties combineren met de kracht van zelflerende technieken. Deze techniek zou op geautomatiseerde wijze mutaties kunnen identificeren die afwijken van het gangbare patroon in de desbetreffende stroom. Hiermee kan het zoekproces verder gestuurd worden en zodoende zouden de (handmatige) taken van accountants verder ondersteund kunnen worden. ■

Q. Rijnders MSc RA studeerde Bedrijfskunde aan de Erasmus Universiteit Rotterdam, gevolgd door een RA-traject bij ABS. Zij werkt als senior manager bij KPMG Forensic Technology met vraagstukken op het snijvlak van data-analyse en finance.

T.A. Eijken MSc heeft Business Mathematics & Informatics aan de VU gestudeerd, is bij KPMG Forensic Technology afgestudeerd en daarna in dienst getreden. De tijdens zijn afstudeerproject ontwikkelde software heeft geleid tot een eerdere artikel in 2011.

M.V.M. Fissette MSc heeft Forensic Science aan de UvA gestudeerd, is bij KPMG Forensic Technology afgestudeerd en daarna in dienst getreden. De tijdens haar afstudeerproject ontwikkelde methode heeft mede geleid tot dit artikel.

J.J.M. van Schijndel MSc heeft Informatica aan de UU gestudeerd, is bij KPMG Forensic Technology afgestudeerd en daarna in dienst getreden. De tijdens zijn afstudeerproject ontwikkelde software heeft mede geleid tot dit artikel.

## Noten

<sup>1</sup> Wij gebruiken de term 'mutatie' voor 'een regel binnen het grootboek (betreft dus geen hele jaartalpost)', omwille van de leesbaarheid van dit artikel.

<sup>2</sup> <http://www.sizes.com/people/brain.htm>

<sup>3</sup> Op internet veelvuldig aangeduid, o.a. op Wikipedia. Een betrouwbare wetenschappelijke

bron hiervoor hebben wij niet gevonden

<sup>4</sup> Handleiding Regelgeving Accountancy, Nader Voorschriften Controle- en Overige standaarden 315.3.

<sup>5</sup> Het Bayesiaanse gedachtegoed maakt het mogelijk om onbekende voorwaardelijke kansen te berekenen door gebruik te maken van beken-

de voorwaardelijke kansen (Russel & Norvig 2002).

<sup>6</sup> Het 'subset sum'-probleem houdt in dat voor een gegeven set getallen (W) en het getal (s), een subset van W (S) gevonden moeten worden waarvan de som gelijk is aan het getal s.

<sup>7</sup> Van de 5 gebruikers in de tweede groep

hebben 4 een RA-titel en 1 heeft een RE-titel.

8 De Journaalposten die uit balans zijn, worden apart terug gerapporteerd.

9 De groepen sluiten zo veel mogelijk aan bij

de captions in de jaarrekening.

10 Te denken valt aan verbandscontroles, gecombineerde statistische toetsing van deelpopulaties binnen het grootboek die consistente pa-

tronen kennen en het benchmarken aan databases met branchedata.

## Literatuur

- Aalst, W.M.P., van der (2011). *Process mining: Discovery, conformance and enhancement of business processes*. Londen: Springer-Verlag.
- Debreceny, R.S., & Gray, G.L. (2010). Data mining journal entries for fraud detection: An exploratory study. *International Journal of Accounting Information Systems*, 11(3), 157-181.
- Haas, M. de (2003). Opinie: De accountant moet 'digitaal'. *De Accountant*, 109, maart, 58.
- Hassink, H. (2002). Accountants en verwachtingskloof: Een onderzoek onder vakbonden. *Maandblad voor Accountancy en Bedrijfseconomie*, 76(7/8), 316-323.
- Keim, D.A. (2002). Information visualization and visual data mining. *IEEE Transactions on visualization and computer graphics*, 7(1), 100-107.
- Keim, D.A., Mansmann, F., Schneidewind, J., & Ziegler, H. (2006). Challenges in visual data analysis. In *Proceedings of the tenth international conference on information visualization* (pp. 9-16). London: IEEE Press.
- Keller, R., Eckers C.M. & Clarkson P.J. (2006). Matrices or node-link diagrams: which visual representation is better for visualising connectivity models? *Information Visualization*, 5, 62-76.
- Koopmans, L. (2005). Real time science ficti-
- on?, *De Accountant*, 111, februari, 22-23.
- Landauer, T. (1986). How much do people remember? Some estimates of the quantity of learned information in long-term memory. *Cognitive Science*, 10(4), 477-493.
- Lari, L.R. (2009). *The power of financial ratios in detecting fraudulent financial reporting: the case of savings and credit co-operative societies in Kenya: the case of Savings and Credit Co-operative Societies in Kenya*. MCom. Thesis, Strathmore University, Nairobi.
- Martello, S., & Toth, P. (1985). Approximation schemes for the subset-sum problem: Survey and experimental analysis. *European Journal of Operational Research*, 22, 56-69.
- Michalewicz, Z., & Fogel, D.B. (2004). *How to solve it: Modern heuristics*. Berlin: Springer-Verlag.
- Mueller-Wickop, N., Schultz, M., Gehrke, N., & Nüttgens, M. (2011). Towards automated financial process auditing: Aggregation and visualization of process models. *EMISA'11* (pp. 121-148). Geraadpleegd op <http://subs.emis.de/LNI/Proceedings/Proceedings190/121.pdf>.
- Rijnders, Q., Özer, P., Blankers, V., & Eijken, T.A. (2011). Zelflerende software detecteert opvallende transacties. Onderzoek naar de mogelijkheden om concepten uit de kunstma-
- tige intelligentie in te zetten voor analyse van financiële gegevens. *Maandblad voor Accountancy en Bedrijfseconomie*, 85(3), 172-181.
- Russell, S.J., & Norvig, P. (2002). *Artificial intelligence: A modern approach* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages*, (pp. 336-343).
- Schouten, R.P. (2007). De toegevoegde waarde van audit software. *Maandblad voor Accountancy en Bedrijfseconomie*, 81(11), 521-530.
- Hoeven, R.L. ter (2005). Principle-based versus Rule-based accounting standards. Waar ligt de balans? *Maandblad voor Accountancy en Bedrijfseconomie*, 79(10), 475-485.
- Veenstra, R.H., & Heertje, A. (2006). Doeltreffende administratieve controle door steekproeven. *Maandblad voor Accountancy en Bedrijfseconomie*, 80(12), 611-619.
- Wallage, Ph. (2005). De actuele waarde van Limpergiaans vertrouwen. *Maandblad voor Accountancy en Bedrijfseconomie*, 79(4), 125-131.
- Zanten, L. van (1892). Het boekhouden als wetenschap. NIVRA Historisch Archief, pp. 10.