

Optimalisatie van Advertenties op Webpagina's



Brian Tromp

BWI-werkstuk

Vrije Universiteit Amsterdam

Faculteit der Exacte Wetenschappen
Vrije Universiteit
De Boelelaan 1083a
1081 HV Amsterdam
Nederland

Begeleider: Ger Koole
2010

Voorwoord

Het BWI-werkstuk is één van de laatste onderdelen van de Master opleiding Business Mathematics and Informatics (BMI). Het doel is een analyse en beschrijving geven van een BMI onderwerp gebaseerd op bestaande literatuur.

In 2010 introduceerde Prof.Dr. Ger Koole mij met “optimalisatie van advertenties op webpagina's”, het onderwerp van dit werkstuk. Aan de hand van het artikel van S. Muthukrishnan (2009) ben ik begonnen me te verdiepen in het onderwerp en heb ik de literatuur bestudeerd. Hieruit kwam naar voren dat veel rendementsverbetering mogelijk is. Nadat ik genoeg kennis had verkregen over het onderwerp heb ik contact gehad met verschillende bedrijven die actief zijn op deze markt. Met de verzamelde gegevens is dit werkstuk tot stand gekomen.

Dit werkstuk is geschreven onder begeleiding van Prof.Dr. Ger Koole van de Vrije Universiteit van Amsterdam. Zonder hem was het niet mogelijk voor mij om dit werkstuk te schrijven. Verder wil ik graag Paul Hissink, Dr. Auke Pot en Dimitri van Esch van Adtrackxys bedanken. Zij hebben mij verder inzicht gegeven in de wereld van online adverteren en hoe hiermee om wordt gegaan. Daarnaast gaat mijn dank ook uit naar Janneke Niessen van Improve Digital.

Samenvatting

Online adverteren kan gezien worden als een interactie tussen drie groepen. Dit zijn de bezoekers, de uitgevers en de adverteerders. Uitgevers bezitten websites die door de bezoekers bezocht worden. Tegelijkertijd adverteren de adverteerders op de website om de aandacht van de bezoeker te trekken.

Een vierde groep die in mindere mate een rol speelt zijn de ad-netwerken. Dit zijn intermediairs tussen adverteerders en uitgevers. Daarnaast is er Real Time Bidding (RTB). RTB is een redelijk nieuw onderdeel in het proces. Dit is een veiling waar een uitgever real time advertentieruimte aanbiedt. Zowel ad-netwerken als adverteerders kunnen hierop bieden.

Het volledige proces van online adverteren ziet er als volgt uit:

1. Een adverteerder wil bepaalde advertenties laten zien op websites. Een adverteerder kan:
 - a. rechtstreeks met een uitgever afspraken maken.
 - b. een ad-netwerk inschakelen.
2. Een bezoeker wil de website van de uitgever bekijken.
3. De uitgever zoekt de beste advertentie bij de bezoeker. Dit kan doormiddel van:
 - a. de campagnes van de adverteerders bij de uitgevers.
 - b. de gemaakte afspraken tussen uitgever en ad-netwerk.
 - c. een veiling via Real Time Bidding.
4. Als via stap 3 geen goede advertentie gevonden wordt, kan de uitgever ervoor kiezen om een huisadvertentie te tonen.
5. De website wordt geladen met de advertentie.
6. Bezoeker ziet de website.

Voor een uitgever zit de grootste uitdaging in stap 3. Een uitgever verdient geld door een betaalde advertentie te laten zien op zijn website. Het ligt aan de gemaakte afspraken hoe de uitgever zijn geld krijgt, maar het komt erop neer dat hoe beter de bezoeker bij de advertentie past, hoe meer geld het oplevert voor de uitgever. In dit werkstuk staat de vraag centraal hoe de uitgever het beste zijn advertentieopbrengsten kan optimaliseren.

Om een goede balans te vinden tussen het verkennen van nieuwe advertenties enerzijds en het benutten van bekende advertenties anderzijds, kan gebruik worden gemaakt van lineair programmeren. Tevens kan rekening worden gehouden met contractuele eisen en voorkeuren. Een nadeel van een lineair programmeringsmodel is de eenvoud waarmee aannames worden gedaan. In de praktijk blijken echter sommige van deze aannames juist het belangrijkste onderdeel te zijn. Ieder gespecialiseerd bedrijf op dit gebied heeft genoeg kennis in huis om het lineair programmeringsmodel op te lossen. Hier kunnen ze niet meer onderscheidend in zijn ten opzichte van elkaar. Het onderscheid wordt gemaakt door de wijze waarop de verschillende bedrijven de historische gegevens weten te gebruiken. Deze historische gegevens kunnen met behulp van data mining omgezet worden in kennis en deze kennis zorgt ervoor dat de beste advertentie aan een bezoeker toegewezen wordt.

Als advies zou ik uitgevers mee willen geven om met een gespecialiseerd bedrijf in zee te gaan om advertentieopbrengsten te optimaliseren. Zij hebben de kennis en middelen om dit effectief en efficiënt te doen.

Inhoudsopgave

Voorwoord	ii
Samenvatting	iii
Inhoudsopgave.....	iv
1 Inleiding.....	1
1.1 Online adverteren	1
1.2 Probleemstelling	3
1.3 Structuur van het werkstuk.....	4
2 Literatuuronderzoek	5
2.1 Probleemformulering.....	5
2.2 Lineair Programmeringmodel	6
2.3 Aanpassing van het Lineair Programmeringmodel.....	7
2.3.1 Clusters.....	7
2.3.2 Ondergrens	7
2.3.3 Contractuele eisen	7
2.4 Voorbeeld.....	9
3 Marktonderzoek.....	12
3.1 Adtrackxys	12
3.2 Improve Digital.....	13
3.3 Admeta.....	14
3.4 AdMeld.....	15
3.5 Doubleclick.....	15
3.6 Overeenkomsten en verschillen tussen de bedrijven.....	16
4 Het verschil tussen theorie en praktijk	17
5 Conclusie	18
5.1 Verder onderzoek	18
6 Literatuurlijst.....	19
7 Bijlage.....	20

1 Inleiding

1.1 Online adverteren

Activiteiten op het internet met betrekking tot online adverteren kunnen gezien worden als een interactie tussen drie groepen. Deze groepen zijn:

- de *bezoekers* die op het internet surfen
- de *uitgevers* die de websites bezitten, controleren, beheren en voorzien van content
- de *adverteerders* die de aandacht vragen van de bezoekers door middel van advertenties op de websites van de uitgevers

Een uitgever verdient geld door advertenties te laten zien op een website. Het is echter niet raadzaam om veel advertenties te plaatsen. Over het algemeen geldt hoe meer advertenties, hoe lager de kwaliteit van de website. Een kwaliteitswebsite trekt meer bezoeker. Hoe meer bezoekers op de website komen, hoe vaker er een advertentie getoond kan worden die geld oplevert voor de uitgever. Een adverteerder is eerder geïnteresseerd in het plaatsen van advertenties op een kwaliteitswebsite. In de praktijk komt het zelden voor dat een website meer dan twee of drie advertenties toont in één beeldscherm.

Uitgevers en adverteerders sluiten contracten af voor een bepaalde periode waarin een advertentie een bepaald aantal keer getoond moet worden tegen een bepaald budget. Dit zijn de zogenoemde campagnes. De meest voorkomende campagnes zijn de bereikcampagnes en de prestatiecampagnes.

Een bereikcampagne is een campagne om zoveel mogelijk aandacht te trekken. Dit kan bedoeld zijn voor nieuwe producten of diensten waarbij de doelgroep nog niet helemaal vast staat of om een merk onder de aandacht te brengen. Veel van dit soort campagnes werken met een Cost per Mille (CPM) eenheid. Dit betekent dat er een bepaald bedrag betaald wordt per duizend weergaves van de advertentie.

Met een prestatiecampagne wil de adverteerder dat de advertentie niet alleen getoond wordt, maar dat er ook op gereageerd wordt. Onder deze reactie wordt meestal een klik op de advertentie verstaan, waardoor de bezoeker naar de website van de adverteerder gedirigeerd wordt. Afhankelijk van wat er afgesproken is in het contract met de uitgever, wordt die betaald naar het aantal clicks (CPC), inschrijvingen (CPL), orders (CPO), of vergelijkbare eenheden.

Een bekende en verwante term aan de hierboven genoemde eenheden is de effectieve cost per mille (eCPM). De eCPM meet de effectiviteit van de verkochte advertentieruimte via CPC, CPL, CPO of anderen. Het vergelijkt de prestatiecampagne-eenheden met de CPM-eenheid.

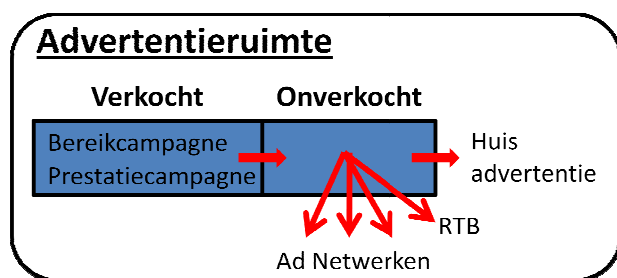
Internet biedt de mogelijkheid om op een website de content en de advertenties te scheiden. Er zijn verschillende manieren om de advertenties te selecteren wanneer er een bezoeker bij een website aankomt. Zo is er de ongerichte, redactionele, gerichte en gepersonaliseerde methode. Onggericht wil zeggen dat de advertenties vast staan of roteren via een vast schema. Deze manier is het makkelijkst maar tevens het minst productief. Redactioneel betekent adverteren op bepaalde websites of onderwerpen. Met de gerichte methode wordt er op een bepaalde voorwaarde gefocust, zoals een type browser of tijdstip of land. Gepersonaliseerd adverteren gaat een stap verder dan de gerichte methode. Hierin worden ook persoonlijke acties mee genomen zoals ingevoerde zoekopdrachten en historisch klik- en surfgedrag.

Grote uitgevers hebben meestal een eigen sales-team om campagnes binnen te halen zodat de advertentieruimte gevuld is. Helaas zijn er vaak niet genoeg campagnes om elk bezoek te beantwoorden met een advertentie. Kleinere uitgevers hebben minder bezoekers maar ook minder mogelijkheden wat adverteerders betreft. Om de advertentieruimte die niet gevuld wordt door

campagnes alsnog te verkopen, kan een uitgever naar een ad-netwerk stappen. Dit zijn intermediairs tussen adverteerders en uitgevers. Een uitgever verkoopt zijn advertentieruimte niet alleen via ad-netwerken omdat de prijs die hij krijgt lager is dan bij advertentieruimte die verkocht wordt via campagnes. Het voordeel voor uitgevers is dat ze makkelijker met meerdere adverteerders in contact kunnen komen waardoor hun onverkochte advertentieruimte alsnog verkocht kan worden. Adverteerders werken via ad-netwerken zodat zij hun campagnes uit handen kunnen geven en zo zelf minder zorgen en kosten hebben om de campagne te leiden. Dit kan met name gelden voor de kleinere adverteerders. Daarnaast staan ze sterker als ze gebundeld zijn in ad-netwerken.

Een redelijk nieuwe techniek is Real Time Bidding (RTB). Dit is een veiling waar advertentieruimtes te koop staan. Iedere adverteerder kan door te bieden een advertentieplaats opkopen en zijn advertentie laten zien. Vanwege de complexiteit wordt deze toepassing nog niet veel gebruikt. RTB is complex om verschillende redenen. Eén reden is dat er meerdere adverteerders en ad-netwerken tegelijk kunnen bieden op dezelfde advertentieruimte. Een andere reden is dat RTB real time gebeurt, waardoor er weinig tijd is om bezoekersinformatie te krijgen, te verwerken en er goed op te reageren. Een derde reden is dat een uitgever het risico loopt op kannibalisatie van zijn advertentieruimte. Met kannibalisatie wordt bedoeld dat adverteerders die normaal gesproken campagnes voeren zich volledig gaan richten op RTB en stoppen met campagnes. Dit is ongunstig voor de uitgever wegens de hogere prijs die gevraagd kan worden voor campagnes ten opzichte van RTB.

Uitgevers kunnen voorwaarden stellen aan alle advertenties die getoond kunnen worden op hun websites. Zo kunnen ze een blocklist samenstellen met advertenties die ze niet willen weergeven en kunnen ze een bodemprijs vaststellen. Bij een blocklist moet men denken aan opdringerige advertenties die de kwaliteit van de website drastisch verlagen. Daarnaast kunnen uitgevers een bodemprijs vaststellen zodat de uitgever een gegarandeerde prijs voor zijn advertentieruimte krijgt. Het nadeel is dat de advertentieruimte mogelijk niet verkocht wordt. Het voordeel is dat er een goede prijs betaald blijft worden voor de advertentieruimte en dat de kwaliteit van de website gewaarborgd blijft. Dit is een afweging tussen korte en lange termijn opbrengsten. In het geval dat de advertentieruimte niet verkocht wordt, kan de uitgever ervoor kiezen om geen advertentie te laten zien of om een huisadvertentie te tonen. Meestal wordt voor de laatste optie gekozen. Huis advertenties zijn advertenties van de uitgever zelf. In Figuur 1 is het proces van de verkoop van advertentieruimte schematisch weergegeven.



Figuur 1

1.2 Probleemstelling

Een uitgever van een website verdient geld als er door een bezoeker op een advertentie geklikt wordt. Het hangt af van welke afspraken er zijn gemaakt met de adverteerder maar in principe geldt: hoe vaker bezoekers op een advertentie klikken, hoe meer opbrengsten het genereert voor de uitgever.

Hieruit blijkt dat de uitgever belang heeft bij een zo goed mogelijke match tussen advertentie en bezoeker. Hoe beter de match, hoe groter de kans dat de bezoeker op de advertentie klikt en daardoor opbrengsten genereert voor de uitgever. Om deze beste match te krijgen, moet er gebruik worden gemaakt van de gepersonaliseerde advertentieselectie methode. Dit is mede mogelijk doordat internet snel genoeg is om voor elk bezoek apart te besluiten welke advertentie er getoond moet worden. Dit gebeurt in de laadtijd van de webpagina. Als dit te lang duurt, verliest de bezoeker zijn interesse en verlaat hij de pagina nog voor deze gezien te hebben. Het is daarom van belang dat de advertentieselectie goed, maar snel uitgevoerd wordt. Wegens de complexiteit van dit proces, zijn er gespecialiseerde bedrijven die dit uit handen van de uitgever nemen en optimaliseren.

Dit werkstuk gaat over het zojuist beschreven proces, gezien vanuit het perspectief van de uitgever. Elke keer als er een bezoeker op de website komt, moet de uitgever uit verschillende advertenties kiezen welke het laat zien. Om een zo groot mogelijke opbrengst te genereren moet de keuze zo gemaakt worden dat de kans het grootst is dat de bezoeker op de advertentie klikt.

Dit kunnen we vertalen naar de volgende probleemstelling:

“Hoe kan een uitgever van websites advertentieopbrengsten optimaliseren?”

Om deze probleemstelling te kunnen beantwoorden splitsen we het op in een theoretische deelvraag en een deelvraag uit de praktijk. Te weten:

“Welke technieken zijn bekend in de literatuur en welke zijn hier van toepassing?”

Er is veel onderzoek gedaan naar rendementsverbetering. Door deze onderzoeken te combineren en toe te spitsen op online adverteren, krijgen we een idee welke techniek hier achter zit en hoe dit gebruikt kan worden.

“Wat is er nu op de markt om rendementsverbeteringen te bereiken voor uitgevers?”

Er zijn verschillende partijen te vinden die dit probleem uit handen van de uitgever nemen. Deze partijen beloven rendementsverbeteringen voor de uitgevers. In dit werkstuk wordt beschreven hoe zij dit doen en wat de voornaamste verschillen zijn tussen deze partijen.

1.3 Structuur van het werkstuk

Het vervolg van het werkstuk is als volgt opgebouwd. In hoofdstuk twee komt de theorie aan bod. We bespreken kort de extra informatie die nodig is om vervolgens het probleem goed te formuleren. Daarna tonen we een model dat een oplossing biedt voor het probleem. Naast dit model wordt een paar aantal aanpassingen gegeven om de werking van het model te verbeteren. Aan het einde van hoofdstuk twee wordt nog een theoretisch voorbeeld gegeven hoe het model werkt.

In hoofdstuk drie worden verschillende bedrijven onder de loep genomen die actief zijn op het gebied van optimalisatie van advertentieopbrengsten voor uitgevers. Achtereenvolgens worden de volgende bedrijven beschreven:

- Adtrackxys
- Improve Digital
- Admeta
- AdMeld
- Doubleclick

In hoofdstuk vier zetten we uiteen wat het grootste verschil is tussen de theorie en de praktijk. Het verkregen theoretische model vergelijken we met algemene punten uit de praktijk. Daarnaast geven we per onderzocht bedrijf kort aan wat de verschillen zijn tussen hun werkwijze en de theorie.

In hoofdstuk vijf worden de conclusies getrokken en wordt antwoord gegeven op de hoofdvraag van het werkstuk. Na de conclusie en aanbeveling is er nog een kleine verwijzing naar een eventueel vervolgonderzoek.

2 Literatuuronderzoek

Relevante onderzoeken over dit onderwerp zijn gelezen en worden in dit gedeelte van het werkstuk behandeld. Wegens de grote diversiteit aan onderzoeken, is er een selectie gemaakt van artikelen waar uitgebreider op in wordt gegaan. Met name de werken van Nakamura et al (2003), Abe et al (1999) en Langheinrich et al (1999) komen in dit hoofdstuk naar voren.

In de teksten van Langheinrich et al (1999) en van Abe et al (1999) wordt er gesproken over het advertentieselectie probleem als een toewijzingsprobleem van een advertentie aan een website. Elke keer dat een bezoeker de website opvraagt, moet er een advertentie aan de website toegewezen worden. Deze teksten veronderstellen dat alle informatie die gebruikt wordt om de beste match te bereiken samen wordt gevoegd tot één array attribuut met alle informatie. Dit array attribuut wordt in het vervolg beschouwd als de input die de bezoeker levert aan de methodes. De informatie die gebruikt wordt, bevat onder andere browserinformatie, informatie van de website en bezoekersinformatie. Browserinformatie bestaat uit bijvoorbeeld zoekwoorden, IP-adres, tijd en datum. Informatie van de website bevat bijvoorbeeld de content, categorieën en sleutelwoorden van de site. Met bezoekersinformatie wordt bijvoorbeeld historisch klik- en surfgedrag bedoeld. Met al deze informatie wordt de beste advertentie uitgezocht. De beste advertentie is die advertentie die de hoogste waarde genereert voor de uitgever. Dit kan afhankelijk zijn van de waarderingseenheid. In de teksten van Nakamura et al (2003), Abe et al (1999) en Langheinrich et al (1999) wordt uitgegaan van de Click-through-rate (CTR) eenheid. Er wordt naar een zo hoog mogelijke CTR gestreefd. De prijs van de CTR is niet relevant voor de methode en wordt daarom niet meegenomen in de teksten. Voor de helderheid en overzichtelijkheid nemen we deze ook in dit werkstuk niet mee

Het toewijzingsprobleem van advertenties op websites op het moment dat een bezoeker aankomt, kan gezien worden als een transportprobleem. Dit probleem kan met lineair programmeren zo optimaal mogelijk opgelost worden. Met een kleine aanpassing is het mogelijk om dit transportprobleem als een 'Hitchcock probleem' te beschouwen.

2.1 Probleemformulering

Stel dat er m banner advertenties zijn, A_1, \dots, A_m , en n array attributen, W_1, \dots, W_n . We gaan er vanuit dat voor elke m advertenties er contracten zijn waarin een minimaal aantal weergaves beloofd wordt over een bepaalde periode, een zogenoemde campagne.

Een campagne bestaat normaal gesproken uit miljoenen weergaves van advertenties. Elke weergave bestaat uit een drietal stappen. De eerste stap is dat één van de n array attributen wordt ingegeven door de bezoeker als input in het systeem. We nemen aan dat de attributen ingegeven worden volgens een vaste kansverdeling. De volgende stap is dat één van de m advertenties aan de website wordt toegewezen. Deze wordt gekozen volgens de huidige kans van weergave. De derde stap is dat de bezoeker de website met advertentie ziet en wel of niet op de advertentie klikt. Dit wordt geregistreerd, waardoor de clickrate via bayesiaanse technieken bijgewerkt wordt en up-to-date blijft. Er wordt vanuit gegaan dat dit proces wordt gegeven door de conditionele kans $P_i(\text{click} | W_i)$, welke vast is maar onbekend.

Het doel is om het aantal clicks te maximaliseren gedurende een campagne terwijl de eisen van het contract worden nageleefd. De uitdaging ligt in de keuze tussen het verkennen en het benutten van combinaties van array attributen en de daarbij toegewezen advertenties. Dit houdt in dat de advertentie per array attribuut een goede kans van weergave moet krijgen. Deze kans van weergave moet in eerste instantie dusdanig hoog zijn dat het model genoeg leert over de werkelijke clickrate van de combinatie. Hierdoor wordt uiteindelijk het aantal clicks gemaximaliseerd, waarbij gebruik

wordt gemaakt van de geschatte clickrates. In dit proces moet het contractuele minimum van het aantal weergaves in acht worden genomen.

2.2 Lineair Programmeringsmodel

Het bovengenoemde probleem kan omgezet worden naar een Lineair Programmeringsmodel (LP-model). Met dit LP-model kan er een oplossing gevonden worden die optimaal is. Het doel van dit LP-model is het aantal clicks te maximaliseren, terwijl het contractueel vastgestelde minimum van het aantal keer dat de advertentie getoond moet worden gerespecteerd wordt. Dit model vinden we terug in de teksten van Nakamura et al (2003), Abe et al (1999) en Langheinrich (1999) en wordt hieronder uiteengezet.

De kans dat er op advertentie A_j geklikt wordt wanneer attribuut W_i opgegeven wordt, is gelijk aan $c_{i,j}$. Neem aan dat de schattingen op deze kansen juist zijn. Neem verder aan dat elk attribuut W_i opgegeven wordt met kans k_i en elke advertentie A_j weergegeven moet worden met kans h_j om te voldoen aan het contractueel vastgestelde minimum. Vervolgens willen we de kans op tonen ($d_{i,j}$) krijgen voor elke advertentie A_j in combinatie met elk attribuut W_i zodanig dat de totale clickrate maximaal is. Dit kunnen we formuleren als het volgende lineaire programmeringsprobleem:

Maximaliseer

$$\sum_{i=1}^n \sum_{j=1}^m c_{i,j} k_i d_{i,j} \quad (1)$$

Onder de condities

$$\sum_{i=1}^n k_i d_{i,j} = h_j \quad (j = 1, \dots, m) \quad (2)$$

$$\sum_{j=1}^m d_{i,j} = 1 \quad (i = 1, \dots, n) \quad (3)$$

$$d_{i,j} \geq 0 \quad (i = 1, \dots, n; j = 1, \dots, m) \quad (4)$$

In woorden komt dit erop neer dat het totaal aantal verwachte clicks gemaximaliseerd wordt. Zie vergelijking (1). Voor elke combinatie van een attribuut met een advertentie, wordt het totaal gesommeerd en gemaximaliseerd. De sommatie wordt pas verhoogd als een bepaald attribuut daadwerkelijk wordt ingegeven, de advertentie daarbij getoond wordt en de bezoeker vervolgens op de advertentie klikt. Dit proces moet plaatsvinden onder de volgende drie voorwaarden. De eerste voorwaarde is dat elke advertentie A_j getoond moet worden met kans h_j . Zie vergelijking (2). Deze kans is de gesommeerde kans over elk attribuut. Deze voorwaarde is nodig om aan de contractuele eisen te voldoen. De tweede en de derde vergelijking volgen uit het feit dat de $d_{i,j}$'s kansen zijn. Wanneer er een attribuut opgegeven wordt, dan is de som van alle kansen per advertentie gelijk aan één. Zie vergelijking (3). Met andere woorden, er wordt altijd een advertentie getoond als er een bezoeker op de website komt. De laatste conditie betekent dat een enkele kans niet kleiner kan zijn dan nul. Zie vergelijking (4).

2.3 Aanpassing van het Lineair Programmeringsmodel

2.3.1 Clusters

Een belangrijke aanpassing op het standaard LP-model is om clusters te maken van de opgegeven array attributen. Abe et al (1999) zegt dit in zijn tekst, net als Nakamura et al (2003). De clusters zorgen ervoor dat er meer data voorhanden is om de clickrates te schatten. Daarbij komt dat het model kleiner wordt, waardoor er minder gerekend hoeft te worden en de methode sneller wordt. Doordat er meer data per cluster is dan per attribuut, worden de schattingen van de clickrates nauwkeuriger. Om hier goed mee om te gaan moet elk attribuut precies onder één cluster vallen en alle clusters bij elkaar moeten alle attributen bevatten. Clusters kunnen verkregen worden door de beste combinatie te vinden van hoge waarschijnlijkheid dat de schattingen van de clickrates juist zijn en de mate van vrije parameters in het model. Hoe meer attributen er in een cluster vallen hoe beter de clickrate te schatten is. Daar staat tegenover dat de clickrate juist verder kan afwijken van individuele attributen. Hoe minder attributen er bij een cluster horen, hoe minder data er is om een schatting te maken van de clickrate. Hierdoor kan het onnauwkeurig zijn. Daar staat tegenover dat de clickrate voor elk individueel attribuut juist beter overeen komt met de clickrate van de cluster.

2.3.2 Ondergrens

De teksten van Langheinrich et al (1999), Abe et al (1999) en Nakamura (2003) wijzen allemaal op een valkuil waar aan gedacht moet worden als het Lineair Programmeringsmodel zijn werk goed doet. Voor een optimale oplossing zal het model veel kansen op weergave op nul zetten. Hierdoor kunnen de combinaties tussen advertenties en attribuutclusters met de hoogste clickrates zoveel mogelijk getoond worden. Echter, het is mogelijk dat sommige combinaties wel een hoge werkelijke clickrate hebben maar ze toevallig in het begin van het proces slecht scoren. Hierdoor is de geschatte clickrate laag en krijgt de combinatie een lage kans op weergave toegewezen. Dit is niet wenselijk wegens de zeer waardevolle potentie van deze combinatie. Dankzij het standaardmodel krijgt het niet de kans om dat te laten zien als het in het begin per toeval fout gaat.

Deze kwestie kunnen we verhelpen door een minimum van de kans op weergave mee te nemen in het model. Deze ondergrens van de kans op weergave kunnen we geleidelijk verlagen, waardoor er in het begin een goed uitgangspunt is voor de verkenning van de clickrate, wat weer tot gevolg heeft, dat er later een reëlere schatting is voor de werkelijke clickrate. Dit betekent voor het model dat vergelijking (4) vervangen wordt door vergelijking (5).

$$d_{i,j} \geq \frac{1}{2m \sqrt{D_{i,j} + 1}} \quad (5)$$

Hier is $D_{i,j}$ het aantal keer dat advertentie A_j getoond is, wanneer cluster W_i is opgegeven. In de vergelijking staat m voor het aantal advertenties. Het moge duidelijk zijn dat hoe vaker deze combinatie weergegeven is, hoe lager de ondergrens van de kans op weergave is. Als de combinatie nog nooit is voorgekomen, is de ondergrens het hoogst.

2.3.3 Contractuele eisen

Een adverteerder kan in de campagne meer eisen opnemen dan alleen het minimum aantal keer dat zijn advertenties weergegeven moet worden. De adverteerder kan bijvoorbeeld aangeven voor welk attribuutcluster hij per se niet wil dat zijn advertentie wordt getoond. De zogenoemde onderdrukkingscombinatie. Een andere mogelijkheid is dat een adverteerder juist wel wil dat zijn advertentie wordt getoond als een bepaalde cluster wordt opgegeven. De adverteerder wil er zeker van zijn dat zijn advertentie met een bepaalde waarschijnlijkheid getoond wordt. De adverteerder

kan de cluster zogenaamd huren voor een bepaald percentage dat de cluster wordt opgegeven door de bezoeker.

Wanneer we te maken hebben met een onderdrukking van een cluster, kunnen we de kans op weergave op nul zetten, waardoor deze combinatie niet voor komt. Dit gebeurt in de tekst van Nakamura et al (2003). Echter, als er veel onderdrukkingen zijn, kan het voorkomen dat er geen oplossing mogelijk is die aan alle eisen voldoet. Wanneer dit het geval is, is er geen controle over de situatie, wat niet wenselijk is. Dit kan verholpen worden door de clickrate van deze combinatie op -1 te zetten. Deze oplossing komt voor in de tekst van Langheinrich et al (1999). Hierdoor kan niet gegarandeerd worden dat de advertentie niet getoond wordt, maar telt dit negatief op bij de maximalisatie functie van het model. Hierdoor wordt deze situatie zoveel mogelijk vermeden, maar als het niet anders kan, is het in ieder geval onder controle.

$$d_{i,j} \geq x$$

(6)

Wanneer de adverteerder een cluster wil huren voor een bepaald percentage, kan vergelijking (4) vervangen worden door vergelijking (6). Hierin is x het percentage waarvoor de adverteerder de combinatie huurt. Deze aanpassing komt zowel voor in de tekst van Langheinrich et al (1999), als in de tekst van Abe et al (1999), en die van Nakamura et al (1999).

2.4 Voorbeeld

In een poging het LP-model te beoordelen wordt deze methode met twee andere methodes vergeleken: een methode die een willekeurige advertentie weergeeft na een opgegeven attribuut en een methode die altijd de advertentie weergeeft, die de maximale clickrate heeft voor het opgegeven attribuut.

De verschillende methodes worden vergeleken aan de hand van een fictief voorbeeld. Dit voorbeeld bevat vijf keuzes uit advertenties die getoond kunnen worden en tien clusters van attributen die opgegeven kunnen worden. De vijf advertenties moeten met respectievelijk fractie 0,40; 0,30; 0,15; 0,10 en 0,05 weergegeven worden om het contractueel minimaal aantal keer van weergave te respecteren. De tien clusters worden uniform ingegeven, wat uitkomt op een kans van 0.10 per cluster. Om de ondergrensvoorwaarde mee te laten doen zodanig dat het nog van belang is, is er gekozen voor een relatief laag aantal keer dat de advertentiecluster combinaties zijn voorgekomen. Deze is voor elke combinatie op 100 gezet. Dit betekent dat voor elke combinatie de kans op weergave minimaal 0,010 moet zijn. De kans op een klik op een advertentie met een gegeven cluster, is willekeurig gekozen tussen de 0,01 en 0,07. In Tabel 6 van Bijlage A staat de precieze toekenning van de kans op een klik voor elke combinatie. Er is in dit voorbeeld gekozen om geen onderdrukking of huurcombinatie mee te nemen. Dit om het voorbeeld simpel en overzichtelijk te houden. De kans op weergave is initieel gezet op 0,20 voor elke combinatie tussen advertentie en cluster. Door de kans op weergave te veranderen, de zogenoemde beslissingsvariabelen, is het de bedoeling een zo groot mogelijke totale clickrate te krijgen.

In Tabel 1 is het resultaat te zien dat verkregen is na oplossing van de verschillende methodes. Hierin is te zien dat de willekeurige selectiemethode het laagst scoort. De lineaire programmeringsmethode verschilt niet veel van de maximale clickratemethode. Echter, de maximale clickratemethode is net iets beter dan de lineaire programmeringsmethode. Als de willekeurige selectiemethode als standaard wordt genomen, zien we dat de lineaire programmeringsmethode bijna 31% beter scoort, terwijl de maximale clickratemethode ruim 33% beter is.

methode	totale clickrate	index
Willekeurige selectie	0,036	1
Lineair Programmering	0,047	1,308
Maximale clickrate	0,048	1,334

Tabel 1

Het verschil tussen de methodes kan beter begrepen worden als er nauwkeuriger naar de kans op weergave wordt gekeken. In Tabel 2 is de kans op weergave te zien voor het willekeurige selectiemodel. Hierin is duidelijk zichtbaar, dat de kansen willekeurig gelijk zijn per cluster en advertentie. De verschillen worden veroorzaakt door de voorwaarde van het contractueel afgesproken minimaal aantal keer van weergave.

In Tabel 3 is de kans op weergave te zien voor het LP-model. Hierin is te zien dat er daadwerkelijk met een ondergrens wordt gewerkt. Geen enkele kans op weergave is kleiner dan de ondergrens, terwijl er voor de combinaties met een grote clickrate een zo groot mogelijke kans op weergave is gerealiseerd.

In Tabel 4 is de kans op weergave te zien voor het maximale clickrate model. Hierin is duidelijk te zien dat er niet met een ondergrens wordt gewerkt. Veel combinaties hebben een kans op weergave van nul. Hierdoor kunnen andere combinaties met een hoge clickrate een kans op weergave van 1 krijgen.

Willekeurige selectie

		Advertentie				
$d(i,j)$		1	2	3	4	5
Array attribuut	1	0,40	0,30	0,15	0,10	0,05
	2	0,40	0,30	0,15	0,10	0,05
	3	0,40	0,30	0,15	0,10	0,05
	4	0,40	0,30	0,15	0,10	0,05
	5	0,40	0,30	0,15	0,10	0,05
	6	0,40	0,30	0,15	0,10	0,05
	7	0,40	0,30	0,15	0,10	0,05
	8	0,40	0,30	0,15	0,10	0,05
	9	0,40	0,30	0,15	0,10	0,05
	10	0,40	0,30	0,15	0,10	0,05

Tabel 2

Lineaire Programmering

		Advertentie				
$d(i,j)$		1	2	3	4	5
Array attribuut	1	0,01	0,16	0,81	0,01	0,01
	2	0,11	0,86	0,01	0,01	0,01
	3	0,96	0,01	0,01	0,01	0,01
	4	0,01	0,01	0,06	0,91	0,01
	5	0,01	0,01	0,56	0,01	0,41
	6	0,01	0,96	0,01	0,01	0,01
	7	0,96	0,01	0,01	0,01	0,01
	8	0,01	0,96	0,01	0,01	0,01
	9	0,96	0,01	0,01	0,01	0,01
	10	0,96	0,01	0,01	0,01	0,01

Tabel 3

Maximale clickrate

		Advertentie				
$d(i,j)$		1	2	3	4	5
Array attribuut	1	0	0	1	0	0
	2	0	1	0	0	0
	3	1	0	0	0	0
	4	0	0	0,16	0,84	0
	5	0	0	0,34	0,16	0,50
	6	0	1	0	0	0
	7	1	0	0	0	0
	8	0	1	0	0	0
	9	1	0	0	0	0
	10	1	0	0	0	0

Tabel 4

In de praktijk zal er vrijwel nooit gebruik worden gemaakt van de willekeurige selectiemethode. Althans, niet als eindmodel. Het model kan wel als vergelijkingsmateriaal dienen waarmee de uitgever kan controleren wat hij heeft bereikt met de optimalisatie. Om alleen met dit model te werken, is het niet goed genoeg. Mocht er geen geld of andere middelen zijn voor uitgevers, dan zal deze nog eerder met een (simpele) ongerichte of redactionele methode werken.

In tegenstelling tot de willekeurige selectiemethode, boekt de maximale clickratemethode wel een goed resultaat. Ondanks dit goede resultaat zal ook de maximale clickratemethode niet veel gebruikt worden in de praktijk. We zien dat veel combinaties een kans op weergave van nul hebben. Dit zijn de advertenties met lage clickrates. De keren dat deze advertenties wel een kans op weergave hebben krijgen zij de slechtste plaatsen, waardoor het eindresultaat het minst negatief beïnvloed wordt. Dit komt doordat er hier alleen een voorwaarde wordt gesteld aan het aantal keren dat de advertentie wordt weergegeven. In de praktijk zal er ook met een schuin oog gekeken worden naar de resultaten van het aantal keren dat de advertentie is weergegeven. Als de advertentie vaak genoeg getoond wordt, maar dit niet het gewenste resultaat oplevert, is de adverteerder alsnog niet tevreden en kan die bij de uitgever weggaan. Dit is niet wenselijk. Ondanks een beter resultaat op korte termijn, is het resultaat op de lange termijn niet positief.

De methode van Lineaire Programmering is hierdoor de beste keuze. Deze behaalt een iets minder resultaat dan de maximale clickrate methode, maar stelt de adverteerders over de hele breedte tevreden, waardoor ook op de lange termijn geen negatieve gevolgen worden ondervonden.

3 Marktonderzoek

In dit deel van het werkstuk komen de specialistische bedrijven naar voren die voor de uitgever het toewijzingsprobleem uit handen nemen. Deze partijen beweren dat zij de beste match tussen bezoeker en adverteerder kunnen maken, waardoor de uitgever het meeste geld krijgt voor zijn advertentieruimte. De volgende bedrijven komen aan bod: Adtrackxys, Improve Digital, Admeta, AdMeld en Doubleclick. Volgens de map van Improve Digital [6] bestaat de omgeving van online adverteerders uit veel meer onderdelen en partijen, maar wegens de scope van dit werkstuk is gekozen om alleen vanuit de opinie van uitgever te kijken. Zelfs in dit afgebakende gebied geldt dat niet alle bedrijven zijn meegenomen. Van al deze bedrijven is echter een zodanige selectie gemaakt, dat alle aspecten aan bod komen.

Advertentieopbrengsten optimaliseren, is met name interessant voor de grotere uitgever. Met grotere uitgever wordt bedoeld dat de website miljoenen unieke bezoekers per dag krijgt. Voorbeelden van zulke uitgever zijn de Telegraaf en RTL.

Net als de advertentieruimte, kunnen ook de gespecialiseerde bedrijven in tweeën gedeeld worden. Het ene deel van de bedrijven specialiseert zich in de optimalisatie van de verkochte advertentieruimte en het andere deel specialiseert zich in de onverkochte advertentieruimte.

Voor de optimalisatie van advertentieruimte zijn de belangrijkste onderdelen de historische data, ad-netwerken en Real Time Bidding (RTB). Deze onderdelen zien we nagenoeg bij ieder bedrijf terugkomen. Daarnaast zijn er ook onderdelen die minder belangrijk zijn of veel overlap hebben met de bekende onderdelen. Denk hierbij aan demanside platform, retargeting, audience buying, ad-exchange en pub-netwerk. In Bijlag B staan de betekenissen van de zojuist genoemde begrippen.

3.1 Adtrackxys



Volgens de CEO van Adtrackxys, Paul Hissink, is Adtrackxys gespecialiseerd in het automatiseren en optimaliseren van verkochte advertentieruimte. Daarnaast hebben ze ook contacten met ad-netwerken om de onverkochte advertentieruimte te optimaliseren. Adtrackxys richt zich niet alleen op rendementsverbeteringen voor uitgever, maar ook voor adverteerders en ad-netwerken. Door deze aanpak hebben ze een brede kennis en begrijpen ze alle kanten van het verhaal.

Adtrackxys onderscheidt zich van hun concurrenten door de optimalisatie van prestatiecampagnes. Daarnaast nemen ze het advertentie management systeem (AMS) van de uitgever niet over, maar integreren zij hun kennis en technologieën in het bestaande AMS van de uitgever. Dit heeft als voordeel dat de uitgever geen grote investeringen hoeft te doen en Adtrackxys zich kan concentreren op de optimalisatie zelf.

Op de website van Adtrackxys [3] is de 'NOAX Ad Sales Optimizer' speciaal bedoeld om rendementsverbetering te behalen voor uitgever. NOAX zorgt ervoor dat de juiste advertentie getoond wordt als de website wordt opgevraagd. De uitgever zorgt hiermee voor de juiste input in het systeem. Volgens Dr. Auke Pot, verzorgt de uitgever ook de informatie en voorwaarden waar NOAX zich aan moet houden. De uitgever heeft de contacten met adverteerders en heeft voldoende kennis om deze door te spelen aan NOAX. Als een adverteerder bijvoorbeeld een specifieke doelgroep wil bereiken, zorgt Adtrackxys hiervoor. Dit realiseren ze door voorwaarden te stellen in

het model. Aan de hand van historische gegevens, kunnen ze aannemelijk maken dat de beoogde doelgroep daadwerkelijk bereikt wordt. Adtrackxys logt alles wat er getoond wordt en hoe er op gereageerd wordt door de bezoeker. De historische gegevens waarover gesproken wordt, vallen onder het beheer van Adtrackxys. Deze logs zijn van grote waarde voor Adtrackxys. Met deze gegevens kan via data mining de juiste kennis worden vergaard. Deze kennis is dan ook de corebusiness van het bedrijf.

Dr. Auke Pot heeft het over de volgende gegevens die van toepassing zijn op de toewijzing van advertenties. Deze gegevens zijn onder andere: type browser, tijd, datum, weer of temperatuur, clickgedrag en geschiedenis van de bezoeker. Daarnaast kijkt Adtrackxys ook nog of de samenhang tussen advertenties elkaar positief of negatief beïnvloedt. Naast de correlatie tussen advertenties onderling, wordt ook gekeken of de content van de website zelf een correlatie heeft met advertenties. Met al deze gegevens 'leert' NOAX de beste toewijzing te maken.

Doordat NOAX bij alle onderdelen van de advertentieruimte betrokken is, kan het een stapje verder gaan dan andere optimalisatiepartijen volgens Paul Hissink. Zo kijkt NOAX zowel wat de beste keuze is voor de bereikcampagnes en prestatiecampagnes, als voor de onverkochte advertentieruimte. Het maakt een afgewogen keuze tussen de onderdelen en in de keuze biedt het de optimale advertentie die de uitgever moet weergeven op de website.

3.2 Improve Digital



Improve Digital is de Europese afdeling van het Amerikaanse PubMatic. Door zich op de Europese markt toe te spitsen kunnen zij op markt specifieke eigenschappen inspelen.

Improve Digital richt zich voornamelijk op uitgevers en niet zozeer op adverteerders volgens COO, Janneke Niessen en de website van Improve Digital [5]. Zij houden zich voornamelijk bezig met de vraag hoe uitgevers hun advertentieruimte kunnen benutten als ze het zelf niet verkocht krijgen. Improve Digital biedt een dashboard om het gehele advertentie management systeem te regelen, maar doet inhoudelijk niet veel met het verkochte deel. Improve Digital biedt een oplossing om zo optimaal mogelijk met onverkochte advertentieruimte om te gaan. Zij stappen in de lege ruimte die de campagnes achterlaten. Zodoende verhogen zij het rendement door in combinatie met historische data gebruik te maken van ad-netwerken en Real Time Bidding (RTB). In werkelijkheid is het percentage dat verkocht wordt via RTB relatief klein. Dit komt mede doordat de ad-netwerken of adverteerders zelf nog niet de functionaliteit bezitten om dit goed te doen. Op de website van Improve Digital [5] staat dat er ook gewerkt wordt met ad-exchanges, demand-side platforms, retargeting, audience buying en pub-netwerken.

Janneke Niessen zegt dat iedere uitgever zijn eigen Ad Network Optimiser (ANO) heeft. Hiermee wordt zoveel mogelijk opbrengst gegenereerd in het onverkochte deel van de advertentieruimte. Elke ANO is anders en kan andere uitkomsten geven. Het grootste gedeelte waarop de beslissing gebaseerd is, is de historische data. Dit hangt echter ook af van de uitgever en de advertentie. Zo kan de ene ANO de historische data van één dag terug meenemen, terwijl een andere ANO de historische data van langer terug meeneemt. Ook geldt over het algemeen dat er meer parameters worden meegegeven in het systeem om de juiste advertentie te kiezen als er een vaste prijs is afgesproken. Om vast te stellen welk ad-netwerk het meeste wil betalen voor een

advertentieruimte, kijkt ANO niet alleen naar het ingegeven attribuut, maar ook naar de minstens net zo belangrijke frequentie van de bezoeker. Zo is een nieuwe bezoeker veel meer waard voor een adverteerder dan een bezoeker die al vaker op een bepaalde site is geweest en de daar geplaatste advertenties al heeft gezien.

Volgens Janneke Niessen werkt Improve Digital veel met blocklists, dit zijn advertenties die niet getoond mogen worden bij bepaalde websites. Zo kunnen ze een hoogwaardig product aan de uitgevers aanbieden. Verder werkt Improve Digital ook veel met een bodemprijs voor advertentieruimte. Als de uitgever deze ruimte zelf niet gevuld krijgt en de ANO ook geen ad-netwerk kan vinden dat minimaal de bodemprijs wil betalen, dan wordt de ruimte opgevuld met huisadvertenties van de uitgever. Het ligt aan de uitgever zelf of ze met een bodemprijs werken of niet. Als er met een bodemprijs gewerkt wordt, bepaalt de uitgever hoe hoog die is.

3.3 Admeta



Op de website van Admeta [2] is te lezen dat Admeta zowel voor uitgevers als voor adverteerders een oplossing biedt om zo goed mogelijk om te gaan met online adverteren. Ze bieden een totaalpakket, wat ze Whitebox TANGO noemen. Dit pakket omvat de verschillende campagnes, een oplossing voor de onverkochte advertentieruimte, een dashboard en een facturatiesysteem. Als de uitgever daar geen behoefte aan heeft, kan ook een afgeslankte versie aan het bestaande AMS toegevoegd worden. Admeta mengt zich niet in het optimaliseren van de verkochte advertentieruimte, maar stapt pas in als het eigen sales-team van de uitgever de advertentieruimte niet verkocht krijgt. Admeta belooft een 100% benutte advertentieruimte terwijl de kwaliteit van de website behouden blijft. Zij zorgen er voor dat de onverkochte advertentieruimte verkocht wordt, zonder risico op kannibalisatie. Dit doet Admeta door verschillende ad-netwerken tegelijk in te zetten om de advertentieruimte te vullen. Daarnaast biedt Admeta de mogelijkheid voor Real Time Bidding. Zo kan een advertentieruimte altijd verkocht worden tegen de beste prijs. Om deze beste prijs te bepalen, heeft Admeta een eigen Ad Network Optimiser. Deze Ad Network Optimiser beslist welk ad-netwerk de advertentieruimte krijgt. Hierbij wordt gekeken naar de hoogste effective cost per mille (eCPM) en niet zozeer naar de hoogste prijs. Het kan zo zijn dat een cost per click (CPC) van 0,8 euro verliest van een 0,5 euro CPC wegens een lagere clickrate. Alles draait hierbij om de eCPM. Dit is het verschil tussen korte termijn en lange termijn opbrengsten.

Het grootste verschil met andere aanbieders van advertentieruimte-optimalisatie is dat beslissingen genomen worden door te concentreren op de browserinformatie en niet op de algemene informatie, dit beweert de website van Admeta [2] althans. De andere parameters worden weliswaar meegenomen, maar de focus ligt op browserniveau. Admeta gebruikt verschillende algoritmes voor de verschillende modellen. Ze laat de keuze aan de uitgever zelf. Deze bepaalt wat hij het prettigst vindt werken. Admeta ondersteunt modellen gebaseerd op CPC, CPM, CPA, CPO, CPL en CPE. Dit zijn zo goed als alle mogelijke waarderingseenheden.

3.4 AdMeld



AdMeld richt zich exclusief op premium uitgevers volgens de website van AdMeld [1]. Dit zijn de grootste uitgevers. Met behulp van AdMeld kunnen zij een 100% bezetting van hun advertentieruimte verwachten zonder aan kwaliteit te verliezen. AdMeld mengt zich niet in de verkochte advertentieruimte, maar stapt pas in als er onverkochte ruimte is. Ze vullen deze ruimte met optimalisatie methodes gebaseerd op historische data. Dit is niet anders dan de meeste vergelijkbare partijen.

De website van AdMeld [1] zegt dat het onderscheidend vermogen van AdMeld voornamelijk in de mogelijkheid ligt om voor meerdere kanalen advertentieopbrengsten te maximaliseren. Met deze kanalen wordt onder andere bedoeld op mobiele ondersteuning. Verder ligt de nadruk op Real Time Bidding. Deze RTB is vergeleken met andere partijen al relatief groot en de verwachting is dat dit alleen nog maar zal toenemen. Desondanks ondersteunt AdMeld ook retargeting, audience buying en ad-exchanges. AdMeld richt zich voornamelijk hierop en wil zich daarmee profileren. Optimalisatietechnieken voor de verkochte advertentieruimte worden niet door AdMeld aangeboden.

3.5 Doubleclick



DoubleClick is de meest volledige van alle partijen die onderzocht zijn. De website van Doubleclick [4] heeft het over een compleet systeem dat wordt aangeboden onder de naam DART for Publishers (DFP). Het bevat alle onderdelen die een uitgever kan gebruiken. Het is aan de uitgever om te bepalen welke onderdelen deze wil hebben. Waar het voor DFP met name om draait is DART Adapt en ad-netwerken. Met DART Adapt biedt Doubleclick optimalisatietechnieken voor zowel de verkochte als de onverkochte advertentieruimte. Voor de onverkochte advertentieruimte kunnen ze gebruik maken van een groot aantal ad-netwerken. Dit komt mede doordat ze de bekendste zijn en onderdeel uitmaken van Google.

Daarnaast ondersteunt Doubleclick onder andere rich media, retargeting, audience buying, mobiele en video ondersteuning en relatie management functionaliteiten. Mocht de uitgever zijn eigen advertentiemanagement systeem hebben, dan heeft Doubleclick tevens een afgeslankte versie die ingebouwd kan worden. Real Time Bidding wordt ook steeds groter en maakt hiermee het plaatje compleet om de onverkochte ruimte te verkopen.

3.6 Overeenkomsten en verschillen tussen de bedrijven

In Tabel 5 staat alles nog een keer in een overzicht en zien we de vijf bedrijven terugkomen met hun specialisaties. Dit is gebaseerd op uitspraken van Auke Pot van Adtrackxys en de websites van de verschillende bedrijven [1], [2], [3], [4] en [5].

Bedrijf	Verkocht advertentieruimte		Onverkochte advertentieruimte	
	Bereik campagne	Prestatie campagne	Ad-netwerk	Real Time Bidding
Adtrackxys	+	++	+	-
Improve Digital	-	-	+	+
Admeta	-	-	+	+
AdMeld	-	-	+	++
DoubleClick	+	+	++	+

Tabel 5

Adtrackxys specialiseert zich op het verkochte deel en doet dat door met name de prestatiecampagne te optimaliseren. Daarnaast hebben ze ook contacten met ad-netwerken om het onverkochte deel te verkopen. Er zijn klanten waarbij ze 75% rendementverbetering bereiken. Improve Digital geeft aan dat zij een verbetering tussen de 60 en 300% realiseren. Dit doen zij door de onverkochte advertentieruimte te optimaliseren. Admeta richt zich ook op het onverkochte deel. Ze zeggen dat ze klanten hebben waarbij ze 200% verbeteren. AdMeld specialiseert zich voornamelijk op Real Time Bidding. Het is niet bekend hoeveel verbetering zij bereiken bij hun klanten. Doubleclick richt zich zowel op verkochte als op onverkochte advertentieruimte. Zij steken wat het aantal ad-netwerken betreft boven de overige bedrijven uit. Doubleclick geeft voorbeelden van klanten waar 37 tot 136% rendementsverbetering is gerealiseerd.

We zien bijvoorbeeld bij Doubleclick dat er een groot verschil zit in de rendementverbetering die zij realiseren. Dit komt door het verschil van optimalisatie in de verkochte en onverkochte advertentieruimte. Als voorbeeld geven ze aan dat een klant met optimalisatie van de verkochte ruimte een verbetering realiseert van 37%, terwijl dat bij een klant die onverkochte advertentieruimte optimaliseert 136% was. Daarnaast zien we ook dat Adtrackxys een lagere rendementverbetering realiseert dan Improve Digital en Admeta. Dit kunnen we, net als bij Doubleclick, koppelen aan het verschil tussen optimalisatie van verkochte advertentieruimte en optimalisatie van onverkochte advertentieruimte.

Improve Digital, Admeta, Admeld en de onderdelen van Doubleclick die de onverkochte advertentieruimte optimaliseren komen erg met elkaar overeen. Adtrackxys en de onderdelen van Doubleclick die de verkochte advertentieruimte optimaliseren komen ook goed overeen.

4 Het verschil tussen theorie en praktijk

In dit hoofdstuk gaan we verder in op de verschillen tussen de onderzochte literatuur en de bevindingen in de praktijk. Ook wordt per individueel bedrijf aangestipt hoe dit met de theorie overeenkomt.

In de teksten van Abe et al (1999), Langheinrich et al (1999) en Nakamura et al (2003) wordt het adverteerprobleem opgelost door middel van een LP-model. In dit model worden aannames en veronderstellingen gedaan. Wanneer deze aannames en veronderstellingen juist zijn, werkt het model uitstekend. Waar het model tekort schiet, is de aanname waarmee informatie verwerkt wordt. Hier wordt bijna geen aandacht aan besteed, terwijl dit juist het belangrijkste deel uit de praktijk blijkt te zijn. Deze omzetting van gegevens in toepasbare kennis is de corebusiness van elk specialistisch bedrijf op het gebied van optimalisatie. Elk onderzocht bedrijf zit op een dusdanig hoog niveau dat het pure toewijzingsprobleem, waar het LP-model bij kan helpen, geen uitdaging meer is. Bedrijven kunnen zich hiermee niet meer onderscheiden van hun concurrenten. Het gaat juist om de verwerking van gegevens, met behulp van data mining, om daarmee de beste keuzes te maken.

Adtrackxys vertoont de meeste overeenkomsten met het LP-model. Ze maken, net als het model, gebruik van array attributen. Het verschil met de theorie is dat Adtrackxys nog een stapje verder gaat door ook rekening te houden met correlaties tussen de advertenties onderling en met de content van de website.

Improve Digital lijkt niet veel op het LP-model. Ze bieden eerder de omstandigheden om de uitgevers zelf zo goed mogelijk de advertentieruimte te laten beheren. Waar Improve Digital veel mee werkt zijn blocklists en bodemprijzen voor advertenties. Dit zijn de onderdelen waar Improve Digital wel lijkt op het LP model.

Admeta komt redelijk overeen met het LP-model. Ze zeggen nadrukkelijk het meest gebruik te maken van browserinformatie. Hiermee geven zij ook aan dat de verwerking van gegevens belangrijker is dan het model zelf.

AdMeld richt zich niet op het LP-model. Zij onderscheiden zich met name door de ondersteuning voor mobiele apparaten en Real Time Bidding.

DoubleClick laat niet veel los over hoe ze rendementverbeteringen bereiken voor de uitgever, maar het is duidelijk dat ze een breed aanbod hebben van mogelijkheden om dit te bereiken. Het kan bijna niet anders dan dat ook Doubleclick gebruik maakt van het LP-model.

Kortom, de onderzochte theorie komt goed overeen met de praktijk. Kanttekening is dat het hier slechts om een gedeelte gaat en niet het gehele proces omvat. De echte moeilijkheid zit hem in de verwerking van de gegevens, de zogenoemde data mining technieken. Hier zien we in de praktijk een oplossing voor iets waar in de theorie niet veel aandacht aan wordt besteed. Als alternatief kan de theorie van data mining vergeleken worden met de praktijk.

5 Conclusie

Het doel van dit werkstuk was om een antwoord te vinden op de probleemstelling: *“Hoe kan een uitgever van websites advertentieopbrengsten optimaliseren?”*. Dit antwoord is verkregen door de praktijk te onderzoeken en te controleren in hoeverre dit overeenkomt met de theorie.

Het LP-model uit de teksten van Abe et al (1999), Langheinrich et al (1999) en Nakamura et al (2003) is in beginsel niet al te moeilijk. Dat betekent echter niet dat de theorie toepassen in de praktijk ook makkelijk is. Het blijkt dat veel grote uitgevers dit uitbesteden aan gespecialiseerde bedrijven. Deze bedrijven hebben zeer veel kennis in huis, waardoor ze effectief en efficiënt rendementsverbeteringen kunnen bereiken voor uitgevers. Deze bedrijven hebben het LP-model wel op orde, maar de grootste uitdaging ligt in de data mining van de historische gegevens. Aan de hand van de hiermee verkregen kennis, is het mogelijk om de beste advertentie toe te wijzen aan een bezoeker. Omdat deze bedrijven dit voor meerdere uitgevers doen, kunnen zij zich hier volledig op concentreren waardoor alles in het teken staat van de optimalisatie.

Als ik een uitgever van websites zou mogen adviseren, zou mijn advies zijn één van deze specialistische bedrijven in de arm te nemen, om zo rendementsverbeteringen te bereiken.

5.1 Verder onderzoek

In dit werkstuk is alleen de gewone advertentieruimte in het onderzoek betrokken. Dit kunnen tekstadvertenties, plaatjes of zelfs video's zijn. Daarnaast kunnen advertenties verschillend zijn van grote en op verschillende posities op de website staan. Dit maakt het hebben van deze vorm van advertentieruimte op de website lastig. Dit is dan ook vaak alleen weggelegd voor grote uitgevers die een eigen salesteam hebben om dit op te pakken.

Naast deze vorm van advertentieruimte, bestaan ook vormen van gerangschikte tekstadvertenties. Dit zijn de zogenoemde AdSense advertenties. Deze vorm van adverteren op websites wordt veelal toegepast bij kleinere uitgevers. Deze hoeven geen sales team te hebben omdat dit een standaard benadering heeft. AdSense gebruikt AdWord wat door Google ontwikkeld is om de beste rangschikking te maken bij zoekopdrachten. Verder onderzoek hiernaar zou nieuwe inzichten kunnen verschaffen over hoe de problematiek op een andere manier opgelost kan worden.

In dit werkstuk is de Real Time Bidding een onderdeel van de optimalisatie van de onverkochte advertentieruimte. We gaan hier niet diep op in aangezien dit buiten de scope van dit werkstuk valt. Aangezien ik verwacht dat RTB in de toekomst een grotere rol gaat spelen dan het nu al doet, is verder onderzoek hiernaar interessant.

6 Literatuurlijst

- Abe, N. et al., *Learning to Optimally Schedule Internet Banner Advertisements*, Proceedings of the 16th International Conference on Machine Learning, p. 3-11, 1999
- Langheinrich, M. et al., *Unintrusive customization techniques for Web advertising*, Computer Networks v.31 p. 1259-1272, 1999
- Muthukrishnan, S., *Ad Exchanges: Research Issues*, Proceedings of the 5th International Conference on Internet and Network Economics p. 1-12, 2009
- Nakamura, A. et al., *Banner Advertisement Selecting Method*, United States Patent, Patent No.: US 6.591.248 B1, 2003

[1] <http://www.admeld.com/publishers.html>, 1-11-2010

[2] <http://www.admeta.com/publishers>, 1-11-2010

[3] <http://www.adtrackxys.com/>, 1-11-2010

[4] <http://www.google.com/doubleclick/publishers/index.html>, 1-11-2010

[5] <http://www.improvedigital.com/>, 1-11-2010

[6] http://www.improvedigital.com/wp-content/uploads/DigitalAdvertisingIndustryMap2010_Netherlands_NL1.0.pdf, 1-11-2010

7 Bijlage

A Clickrate van het voorbeeld

		Clickrate				
		Advertentie				
Array attribuut	c(x,j)	1	2	3	4	5
	1	0,02	0,03	0,04	0,02	0,03
	2	0,03	0,04	0,04	0,02	0,04
	3	0,04	0,04	0,02	0,06	0,05
	4	0,02	0,02	0,06	0,06	0,05
	5	0,01	0,02	0,04	0,04	0,07
	6	0,02	0,04	0,02	0,03	0,04
	7	0,06	0,06	0,05	0,02	0,05
	8	0,04	0,06	0,01	0,05	0,05
	9	0,05	0,04	0,05	0,05	0,05
	10	0,04	0,01	0,06	0,05	0,06

Tabel 6

B Begrippen

Ad-exchange

Een ad-exchange is een online marktplaats voor het kopen en verkopen van advertenties.

Demandside platform

Een demandside platform is een gestandaardiseerde platform voor adverteerders om advertentieruimte te kopen.

Retargeting

Retargeting is het adverteren op bezoekers die al bekend zijn in het systeem. Er is van deze bezoeker bekend hoe die op bepaalde advertenties heeft gereageerd en in welke onderwerpen hij geïnteresseerd is.

Audience buying

Audience buying is het aankopen van informatie van een bepaalde doelgroep. Met deze informatie, van meestal derde partijen, kan bijna met zekerheid de doelgroep bereikt worden.

Pub-netwerk

Een Pub-netwerk is een uitgevers netwerk. Hierin zijn uitgevers gebundeld en worden vertegenwoordigd door het netwerk.

Rich media

Rich media zijn advertenties die uniek zijn in zijn soort en maat. Rich Media wordt vaak op maat gemaakt voor een adverteerder voor een speciale uitgever.