Bachelorscriptie i.o.v. Radboud Universiteit Nijmegen

Profiling onder studenten Informatica en Informatiekunde

Rogier Lommers

14 juni 2006

Samenvatting

Kennis is macht, wordt er vaak gezegd. Met de komst van de informatiemaatschappij en dus de invoering van duizenden met elkaar verbonden computers wordt er dagelijks enorme hoeveelheden informatie uitgewisseld van computer tot computer, van land tot land en zelfs van continent tot continent. Mensen maken gebruik van geavanceerde zoeksystemen om snel en eenvoudig informatie te vinden. Door deze zoekwoorden op te slaan en te hergebruiken, is het mogelijk om als exploitant van een zoekmachine allerlei gebruikersprofielen te creren. Dit heet profilen. De informatie afkomstig van de gebruikers van de zoekmachine wordt gebruikt voor andere doeleinden dan voor de zoekmachine alleen. Wat zijn de mogelijkheden van profilen en wat vinden de internetgebruikers hier van? Is het eigenlijk wel verantwoord dat zoekmachines deze informatie gebruiken?

Inhoudsopgave

1 Onderzoeksplan

1.1 Inleiding

Als je als bedrijf informatie verzamelt van de mensen die je dienst gebruiken, dan moet je er rekening mee houden dat je zorgvuldig met deze informatie om gaat. Neem nu een bedrijf als Google[1]. Elke dag maken miljoenen mensen gebruik van de primaire dienst van Google: de zoekmachine. Door nu deze zoekopdrachten (de querys) te groeperen en te koppelen aan een individueel profiel, kan Google erg veel over mensen te weten te komen. Ook is het mogelijk om gebruik te maken van informatie, afkomstig van gebruikers, om nieuwe informatie te genereren of om deze te verkopen. Voorbeelden zijn de interessegebieden van jeugdige gebruikers of de verschillen tussen man en vrouw. Deze nieuwe gegenereerde informatie kan op verschillende manieren worden gebruikt. Met mijn bachelorscriptie wil ik gaan onderzoeken hoe de informatica- en informatiekunde studenten van de beta faculteit van de Radboud Universiteit Nijmegen tegen dit 'profilen' aankijken.

1.2 Probleemstelling

Tijdens mijn bachelorscriptie wil ik de volgende onderzoeksvraag gaan beantwoorden:

> *Welk beeld hebben studenten informatica en informatiekunde van de Radboud Universiteit Nijmegen van de mogelijkheden en onmogelijkheden van profilen?*

Als product van dit onderzoek wil ik het antwoord op de onderzoeksvraag opleveren waarin staat beschreven hoe het beeld dat bovengenoemde studenten van de Radboud Universiteit in Nijmegen hebben eruit ziet. Wat denken zij dat er tegenwoordig allemaal mogelijk is en toegepast wordt?

1.3 Verantwoording

Tegenwoordig kunnen internetexploitanten erg veel bereiken door het bovengenoemde profilen toe te passen. Een simpele zoekmachine levert al genoeg input om uitgebreide lijsten te genereren die waardevol kunnen zijn voor hergebruik. Een actueel punt is de recente wetswijziging. Op 14 december jl. is het Europees Parlement akkoord gegaan met de bewaarplicht van communicatiegegevens. Vanaf juli 2007 moeten alle lidstaten gegevens van telefoon-, email en internetverkeer voor minimaal zes en maximaal 24 maanden opslaan. Deze ontwikkelingen hebben een verschuiving in de machtsverhoudingen teweeggebracht. Voorheen kon men redelijk anoniem over het net surfen en werden er geen logbestanden bijgehouden. Dit gaat dus drastisch veranderen en maakt het interessant om te bestuderen hoe bepaalde groepen zich hiervan be-

[1] http://www.google.com

wust zijn. De kans is groot dat deze groepen mensen hun internetgedrag aanpassen aan deze wetswijzigingen om zo de risico's van profiling te verkleinen.

1.4 Theoretisch kader

Het onderzoek richt zich op informatiekundig gebied. Er komen zowel technische (de techniek achter het profilen) als sociale aspecten (welke doelgroepen zijn er op de hoogte van profilen en waarom) aan de orde. In hoofdstuk 3.4 worden de verschillende soorten van tegengaan van profilen in kaart gebracht. De manieren waarop het mogelijk is om met profilen om te gaan worden verdeeld in zes privacy-lagen[4]. De reden voor het gebruik van dit zes-lagenmodel is dat het bijdraagt aan het beantwoorden van de onderzoeksvraag; het model maakt het namelijk mogelijk om het bewustzijn van profilen van een individu in kaart te brengen. Met andere woorden: internetgebruikers kunnen worden onderverdeeld in categorieen en met elkaar worden vergeleken. Vervolgens verklein je de kans op misbruik van profiling, aangezien er een maatstaaf wordt gecreëerd waarin internetgebruikers de bedreigingen op persoonlijk niveau ervaren en tegengaan door anti-profilingmiddelen te kiezen, behorende bij hun eigen categorie. De volgende zes lagen zijn aanwezig:

1. awareness
2. control
3. privacy-enhancing tools
4. privacy policies
5. privacy and trust certification
6. privacy-protection laws

De lagen lopen in chronologische volgorden op, van niet-bekend zijn met profiling tot zeer bekend. Dit kader staat centraal tijdens deze scriptie; het wordt gebruikt als referentiekader om de doelgroep in te plaatsen.

1.5 Methode

Om te onderzoeken hoe het beeld dat de studenten van profiling hebben eruit ziet ga ik eerst in kaart brengen welke vormen van profiling er allemaal zijn. Dit als een soort van vooronderzoek om de hoofdvraag (onderzoeksvraag) te kunnen beantwoorden. Ik ga me verdiepen in de technieken achter het profilen, met de bijbehorende voor- en nadelen. Vervolgens ga ik de Radboud studenten benaderen door middel van een online enquête. Ik heb voor een online enquête gekozen, omdat ik van mening ben dat juist deze groep studenten regelmatig op het internet te vinden is. Voordat de enquête wordt afgenomen wil ik de aandacht van de studenten trekken door ze twee situaties voor te leggen waarin profiling toegepast wordt. Op deze manier wil ik proberen snel te laten zien wat

ik bedoel met het onderwerp profiling en om ze te laten focussen op de materie. De studenten moeten met een duidelijk beeld beginnen aan de enquête. De deelvragen helpen mij met het beantwoorden van de onderzoeksvraag. Deze vragen kunnen als volgt worden omschreven:

- Wat is profiling precies?
- Welke soorten van profiling zijn bekend en wat zijn de verschillen?
- Welk soort profiling wordt het meest toegepast?
- Welke doelgroepen komen in aanmerking voor profiling en welke doelgroep is het meest geliefd?

Deze hierboven genoemde zes privacylagen gebruik ik als basis voor dit onderzoek. Ik wil dan ook gaan onderzoeken hoe de studenten informatica en informatiekunde van de faculteit NIII geplaatst kunnen worden binnen deze zes lagen.

1.6 Tijd- en faseringsschema

Het volgende schema geeft aan hoe de planning eruit ziet.

Week	Onderdeel
40	Goedkeuring onderzoeksplan
40	Goedkeuring onderzoeksplan
41	Literatuur zoeken
42	Literatuur bestuderen
43	Literatuur bestuderen
44	Materiaal verzamelen
45	Materiaal verzamelen
46	Materiaal verzamelen
47	Materiaal verzamelen
48	Materiaal verzamelen
49	Materiaal verzamelen
50	Materiaal verzamelen
51	Verwerken verkregen materiaal
52	Verwerken verkregen materiaal
01	Verwerken verkregen materiaal
02	Rapporteren
03	Rapporteren
04	Rapporteren
05	Afronden bachelorscriptie

1.7 Literatuur

Om tot een succesvol antwoord op de onderzoeksvraag te komen is het noodzakelijk dat ik mezelf inlees in de materie. In overleg met dhr. L. Consoli[2] wordt er nog bepaald welke literatuur vereist is om deze bachelor scriptie af te ronden.

[2] http://www.ru.nl/fil-beta/lucac/

2 Inleiding

Het internet zoals we dat nu kennen, wordt alsmaar groter. Was het in 1969 oorspronkelijk ontworpen als een communicatie tool voor het Amerikaanse leger, tegenwoordig is bijna alles online mogelijk. Inmiddels is het internet een wereldomvattend fenomeen dat het karakter van een massamedium heeft gekregen. Het afsluiten van een hypotheek, het kopen van cd's en het snel even opzoeken van iemands telefoonnummer, het is allemaal mogelijk op het World Wide Web. Maar bij alle handelingen die verricht worden, is het noodzakelijk dat er gegevens (digitaal) over de verbinding worden verstuurd naar de eigenaar van de desbetreffende website. Wat wordt er met deze gegevens gedaan? Het koopgedrag van de klant ligt er als het ware mee op straat. Men kan precies zien wat er gekocht wordt, door wie er gekocht wordt en wanneer de aankoop verricht is. Op deze manier is het voor deze zoekmachines mogelijk om voor iedereen een persoonlijk profiel bij te houden met deze gegevens. Ook kunnen deze gegevens worden gebruikt om de resultaten aan te passen aan de gebruiker. In het ergste geval worden de persoonlijke gegevens verkocht aan derden. Met deze scriptie wil ik duidelijk maken wat de mogelijkheden zijn van profilen en hoe studenten van opleidingen Informatica en Informatiekunde van de Radboud Universiteit Nijmegen hier tegenaan kijken.

Om te beginnen ga ik uitleggen met wat profilen precies is. Ik begin met het identificeren van het probleem. Het zogenaamde profilen wordt besproken. Vervolgens beschrijf ik de diverse soorten van profilen; wat de mogelijkheden zijn en hoe internetexploitanten omgaan met de verkregen informatie. Vervolgens ga ik onderzoeken hoe de doelgroep omgaat met het feit dat er informatie wordt gebruikt voor profile-doeleinden. Dit gebeurt door middel van het afnemen van online enqutes. Nadat deze afgenomen zijn ga ik de verkregen gegevens analyseren en trek ik conclusies.

Veel leesplezier toegewenst,

Rogier Lommers 21 december 2005

3 Profiling

3.1 Internetbedrijven

Door de enorme groei van het internet lopen gebruikers het risico om de dupe te worden van profiling. Het risico is het grootst bij websites die erop gericht zijn om de klant persoonlijke of financiële producten te verkopen, of om een online dienst aan te bieden. Deze sites worden business-to-consumer (B2C) sites genoemd. Dit soort websites kunnen herkend worden door middel van enkele eigenschappen[2], te weten:

- Zoeksystemen. Door middel van een inputveld op de website wordt de mogelijkheid geboden om naar bepaalde trefwoorden te zoeken. Vaak is het ook mogelijk om door middel van logische operatoren (AND, OR) de zoekopdracht te verfijnen.

- Multimedia-enabled interactie tussen server en gebruiker. Duidelijke animaties laten de gebruiker zien wat de mogelijkheden van de website zijn. Een simpele muisklik is meestal voldoende om direct naar een bepaalde productcategorie te gaan.

- Tijdsgebonden en persoonlijke inhoud (content). De inhoud van de website verschilt van tijd tot tijd. Zo zijn er tijdens de feestdagen themas/kleuren anders dan tijdens de zomervakantie. Ook is het mogelijk om de website te voorzien van een eigen inhoud. Door middel van het aanmaken van een gebruikersaccount wordt deze persoonlijke inhoud vastgelegd.

- Mogelijkheden om producten te vergelijken. Bij het online verkopen van producten wordt vaak een dienst aangeboden om twee of meerdere producten met elkaar te vergelijken op kwaliteit, prijs en mogelijkheden.

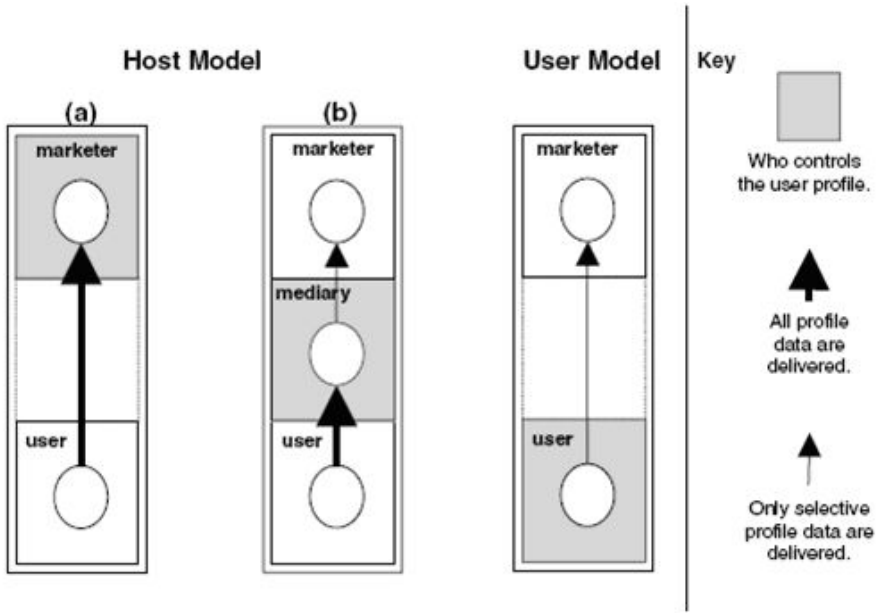
Wat zijn nu precies de mogelijkheden van de internetbedrijven om om te gaan met de persoonlijke gegevens? Volgens het artikel[1] zijn er tegenwoordig twee soorten internetbedrijven actief, namelijk:

- Internet Marketers

- Internet Mediaries

Zie figuur 1 voor een overzicht van waar de gebruikersprofielen bij beide soorten bedrijven opgeslagen worden. De eerste bedrijfstak heeft de naam internet marketers. Dit zijn organisaties die geld verdienen door het verkopen van goederen en services. Een typisch voorbeeld van een Internet Marketer is Bol.com [http://www.bol.com]. Bol is een groot internationaal bedrijf dat boeken, cdroms, dvds, muziek en elektronica artikelen online verkoopt. De tweede soort organisaties, de Internet Mediaries, bieden diensten aan om zo hun geld te verdienen. Voorbeelden van deze diensten zijn internetverbindingen, zoekmachines, startpaginas of email functionaliteiten. Google [http://www.google.com] is een

voorbeeld van een bedrijf uit de klasse Internet Mediaries. Google levert haar klanten verschillende soorten diensten, variërend van email en routeplanners tot de meest bekende dienst: de zoekmachine. Hiermee is het mogelijk om binnen enkele seconden negen miljard websites te doorzoeken.



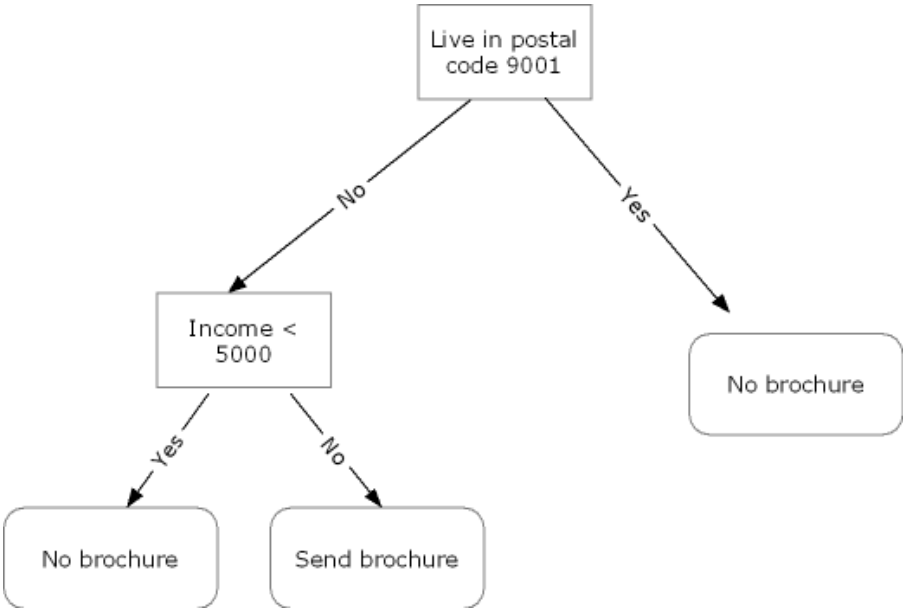
Figuur 1: Locaties van profielen, bron: Spiekermann [1]

Beide soorten bedrijven hebben elke dag te maken met bergen informatie van alle internetgebruikers die gebruik maken van hun product of dienst. Vaak is deze informatie persoonlijk, bijvoorbeeld wat de woorden zijn waarop gezocht wordt bij Google. De auteurs van het artikel[1] hebben negen internetbedrijven bestudeerd. Ze bekeken wat deze bedrijven doen met de door de bezoekers gegenereerde informatie en op welke manier hier gebruik van gemaakt wordt.

3.2 Methoden van gebruik van klantgegevens

Het gebruik van de klantgegevens kan in twee categorien worden verdeeld, namelijk intern en extern. Intern gebruik houdt in dat het bedrijf, bijvoorbeeld Google, de gegevens gebruikt om hun eigen producten of diensten te verbeteren om zo de klant beter van dienst te kunnen zijn. Klantinformatie wordt op deze manier door internetbedrijven gebruikt om de kwaliteit en de effectiviteit van het bedrijf en het product/dienst te verbeteren. Volgens het artikel passen vier van de vijf marketers de zogenaamde service differentiation toe binnen het interne gebruik van informatie. Service differentiation houdt in dat de klanten worden verdeeld in verschillende categorieën, bijvoorbeeld A, B, C en D. Aan de hand van in het verleden gekochte goederen, demografische gegevens of surfgedrag wordt de bezoeker in een bepaalde categorie geplaatst. Zie figuur 2 voor een voorbeeld van service differentiation. Gebruikers met postcode 12101

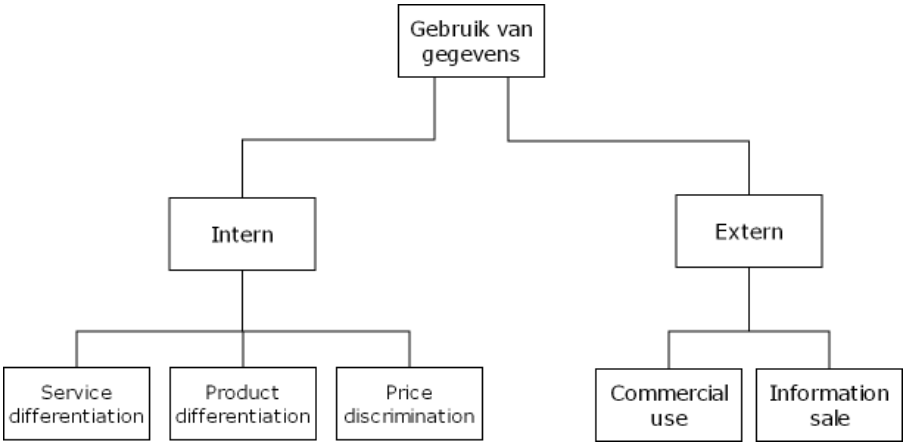
worden anders behandeld dan gebruikers met andere postcodes. Vervolgens wordt er ook nog een verschil gemaakt tussen het inkomen; bij een inkomen dat kleiner is dan 5000 euro wordt er geen reclamefolder verstuurd, bij een hoger inkomen wel. Zo zijn er veel mogelijkheden met de categorieën. De inhoud van de verschillende categorieën worden dan allemaal anders gebruikt. De ene categorie gebruikers bestaat bijvoorbeeld uit klanten die veel kleine artikelen kopen. Reclameboodschappen zoals hierboven worden dan ook speciaal voor deze groep mensen aangepast zodat er gerichte reclame wordt gemaakt; kleinere, goedkopere artikelen gelden dan als aanbieding.



Figuur 2: Filtering op demografische ligging, bron: R.H.W. Lommers

Ook wordt er vaak gebruik gemaakt van product differentiation. Deze methode bestaat uit twee deelmethodes, namelijk up-selling en cross-selling. De eerste methode houdt in dat de internetonderneming met de informatie die ze van de klanten hebben andere producten aanraden die nét iets beter zijn, maar daarvoor moeten klanten natuurlijk ook nét iets dieper in de buidel tasten. Cross-selling wil zeggen dat er aan de hand van de opgeslagen informatie van de klant wordt bepaald welke producten er nog meer in aanraking komen om verkocht te worden aan de klant. Als iemand bijvoorbeeld veel producten koopt zoals MP3-spelers en andere gadgets, dan wordt hierop ingespeeld door de reclame persoonlijk te maken. Er zal gekozen worden om andere gadgets als reclame banner aan te bieden. Ook wordt er, volgens de studie van de auteurs van het artikel, door vier van de vijf marketers veel gebruik gemaakt van de product differentiation methodes up- en cross selling. Veelgebruikte methodes zijn het personaliseren van de emaillijsten. Door ‘persoonlijke‘ mails te sturen naar de verschillende categorieën van gebruikers zorg je er voor dat deze mail als interessant wordt bestempeld door de ontvanger. Door deze persoonlijke benadering verhoog je de omzet. ‘One of our interviewees, with several million

clients, claimed to offer special products and services to segments of as small as 10.000 recipients with similar profiles. As a result, the company profits from a return on marketing investment around three times higher than if it had contacted all its clients indiscriminately.'[1], pagina 700.



Figuur 3: Overzicht profilingtechnieken, bron: R.H.W. Lommers

Het extern gebruik van gegevens houdt in dat het surfgedrag van de gebruiker dermate wordt opgeslagen dat er erg veel bekend is. Zo is er het bedrijf Doubleclick [http://www.doubleclick.com/us] dat centrale reclameboodschappen registreert. Websites plaatsen via de webserver van Doubleclick een reclameboodschap op hun site. De gebruiker klikt er vervolgens op en Doubleclick registreert de gegevens. Op deze manier wordt het gedrag van de surfen vanaf meerdere websites (hosts) geregistreerd en is het mogelijk om deze verschillende informatie te linken. Op deze manier wordt er een uitgebreid profiel van de internetter gecreëerd met de nodige gevolgen van dien. Deze manier van gegevensverzameling valt onder de noemer targeting advertisements.

Een andere manier van extern gebruik van gegevens is het verkopen of verhuren van het internetgedrag van gebruikers. Veel mensen zien echter niet in dat de gegevens waar het hier om gaat nogal uitgebreid zijn. Soms gaat het alleen om gegevens als naam en email adres, maar er zijn ook gevallen waar er veel meer persoonlijke informatie bekend is. Denk hierbij aan NAW gegevens, sofi-nummers en de adressen van de door bezoekers bezochte websites; dus de persoonlijke interessegebieden. Zie figuur 3 voor een overzicht van de soorten profiling-technieken.

3.3 Datamining

De basis van profiling draagt de naam datamining. Datamining is een techniek waarbij getracht wordt om op een geautomatiseerde manier patronen en relaties te ontdekken in grote hoeveelheden gegevens [16]. Deze gegevens kunnen afkomstig zijn van een eigen informatiesystemen, zoals bij SAP of andere

ERP-systemen [3] of van grote webshops als Bol.com. Net zoals bij profiling levert datamining nieuwe informatie op. Echter ook de problemen van profiling zoals privacy, legaliteit en ethiek zijn bekend bij datamining. De onderdelen van datamining kunnen als volgt worden omschreven. [15]

- Dependency analysis. Dit is de meest gebruikte vorm van datamining. Het houdt in dat er geautomatiseerd gezocht wordt naar twee soorten informatie die met elkaar matchen. Als iemand bijvoorbeeld gezocht heeft naar het woord 'saus', dan is er een regel die ervoor zorgt dat er als resultaten ook de producten als 'mosterd' of 'curry' verschijnen.

- Class identification. Deze vorm zorgt ervoor dat bezoekers in vooraf gedefinieerde klassen worden verdeeld. Dit zijn klassen met eenheden (mensen) van bijvoorbeeld hetzelfde inkomen of met gelijke, in het verleden gekochte goederen.

- Concept description. Concept description is een methode om door middel van het groeperen van klanten en hun recentelijke koopgedrag nieuwe concepten te creeren. Een voorbeeld is een productlijn van de 'extra functies' bij een nieuw aan te schaffen auto. Met gegevens van in het verleden gekochte auto's wordt een totaalpakket samengesteld. Vaak wordt concept description gecombineerd met class identification.

- Deviation detection. Met deze techniek worden combinaties gezocht die afwijken van het normale, verwachte gedrag van een systeem. Een voorbeeld is het vergelijken van twee groepen creditcard klanten. De ene groep heeft in een bepaalde periode veel meer uitgegeven dan de andere. Deze verhoging kan als oorzaak hebben dat inkomens zijn verhoogd en deze 'kennis' wordt automatisch opgenomen in het systeem. Het systeem is zelf-lerend.

- Data visualisation. Deze techniek maakt het mogelijk om geavanceerde koppelingen op een 3d-weergave en in kleur te laten zien. Het grote voordeel hievan is dat pieken en dalen veel sneller worden gevonden dan wanneer deze gegevens enkel als tekst worden bekeken.

3.4 Tegengaan van profiling

Het tegengaan van het verzamelen van persoonlijke gegevens is moeilijk te realiseren. In de loop van de jaren zijn er verschillende platformen opgericht die pleiten voor anoniem en niet-persoonsgebonden surfgedrag. Deze platformen hebben technieken ontwikkeld waardoor deze idealen nageleefd kunnen worden en dragen de naam 'PET'. De afkorting PET staat voor Privacy Enhancing Technologies en er zijn vijf soorten. [10]

[3] Enterprise Resource Planning-systemen, grote, brede applicaties die alle administratieve lagen omvatten.

- Encryption and stenography ([11]Sellars, 1999): gegevens worden verborgen gehouden door ze te converteren in een onleesbaar formaat voor kwaadwillenden. Tevens worden de gegevens soms in andere dummy-gegevens verstopt, zodat ze alsnog onopgemerkt en anoniem blijven. Door uitgebreide ontsleutelingstechnieken is het echter onmogelijk om 100 procent te garanderen dat deze methode waterdicht is.

- Blind digital signatures ([12]Senicar et. Al, 2003): door gegevens digitaal te ondertekenen en hierbij niet je naam te gebruiken blijft de identiteit van de gebruiker verborgen; alleen de publieke handtekening is bekend bij de kwaadwillenden. Er is dus met zekerheid te zeggen dat deze gegevens afkomstig zijn van een bepaald persoon.

- Digital cash ([13]Driscoll et al, 1997): Anoniem betalingsverkeer wordt hiermee mogelijk. Het is te vergelijken met de chip-card. Virtueel geld wordt op een vertrouwelijke website in een account gezet, waar vervolgens mee betaald kan worden. Ook hierdoor blijven klantgegevens verborgen.

- Trust centers: bedrijven worden aangesteld om als vertrouwelijke partner op te treden. Deze bedrijven zorgen vervolgens voor de uitgave van persoonlijke certificaten en zorgen voor de verbindingsbrug tussen individuele personen en hun certificaten.

- Anonymizers: ([14]Claessens et al, 1999): methoden om privacy te waarborgen. Voorbeelden zijn cryptografie, pseudoniemen en Proxy servers. Deze vorm van PET wordt anno 2006 redelijk veel gebruikt door de gemiddelde internetgebruiker. Het is bijvoorbeeld mogelijk om anoniem over het World Wide Web te surfen door het bedrijf Anonymizer [http://www.anonymizer.com] te raadplegen. Dit is een bedrijf dat een gateway-dienst bezit. Elke internetgebruiker heeft namelijk een uniek nummer: het ip-adres. Dit is een nummer dat door providers toegewezen wordt aan alle computers die verbonden zijn met het internet. Door gebruik te maken van Anonymizer wordt het echte ipadres verborgen en surf je met het ip-adres van Anonymizer. Het is dus niet meer mogelijk om het prive-adres van de gebruiker op te slaan. Ook zijn er speciale softwarepakketten die er voor zorgen dat de pakketjes data met persoonlijke informatie niet verzonden worden naar de marketers of mediaries. Een voorbeeld van een dergelijk programma is WebWasher[4] [http://www.webwasher.com]. Op deze manier wordt het verzamelen van persoonlijke gegevens tegen gegaan en kan er redelijk anoniem over het internet gesurft worden. Door onwetendheid van de mogelijkheden van profilen maken slechts weinig mensen gebruik van diensten als Anonymzer.com en WebWasher. Veel gebruikers weten wel dat websites persoonlijke informatie verzamelen, maar slechts weinig gebruikers kennen de consequenties van het verzamelen. De zogenaamde privacy policies, de stukken tekst die de gebruikers vertellen wat er precies gedaan wordt met

[4]Het blijft altijd mogelijk om op de servers van Webwasher het IP-adres van de eindgebruiker te achterhalen.

de gegevens, zijn vaak in technisch taalgebruik geschreven en gebruikers snappen er niets van. Andere nadelen zijn het verdwijnen van de voordelen van profiling en de kosten die verbonden zijn aan het gebruik van deze diensten. Naast bovengenoemde anonymizers zijn er ook andere bedrijven die diensten aanbieden om profiling tegen te gaan. Feit blijft echter wel dat de publieke ip-adressen van de gebruikers wel bekend zijn bij deze anonymizers. Indien zij dus gedwongen worden, door bijvoorbeeld politieke overmacht, dan is het alsnog mogelijk dat de publieke ip-adressen bekend worden en alsnog de echte identiteit aan het licht komt. Erger nog: een complete lijst met alle activiteiten van een bepaalde tijdsinterval van een ip-adres is dan ineens beschikbaar.

De kracht achter de bovengenoemde PET-technologieen rust dus in het feit dat ze, in principe, de echte identiteit van de gebruiker geheim houden of gegevens op een dermate wijze versleutelen dat ze onbruikbaar zijn als de juiste sleutel niet bekend is bij kwaadwillenden.

3.5 Privacy in lagen

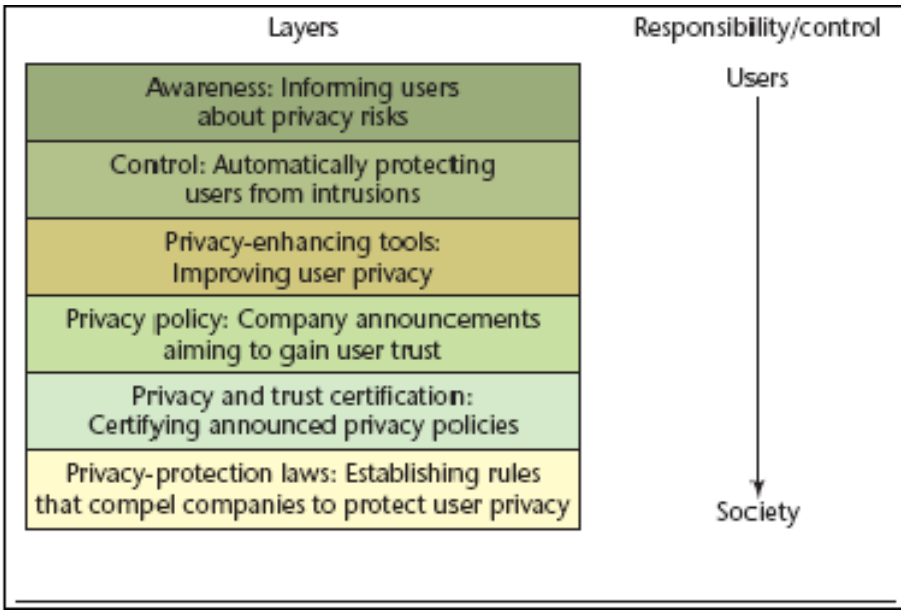
Om internetgebruikers bewust te maken van de gevaren van profiling en om ze in te laten zien in hoeverre ze beschermd zijn, zou je de methoden om de privacy te beschermen in lagen kunnen verdelen. Door lagen te gebruiken stel je gebruikers in staat om het bewustzijn te vergroten en hiermee dus ook de veiligheid. Net zoals geografische lagen zijn ook deze lagen afzonderlijk van elkaar. Als een gebruiker een bepaalde laag, bijvoorbeeld de laag ‘Privacy-enhancing tools: improving user privacy‘, in gebruik heeft, dan wil dat niet zeggen dat hij/zij ook de voorgaande lagen in gebruik heeft. Op de volgende paragrafen bespreek ik de zes lagen en leg ik kort uit wat elke laag inhoudt[4].

3.5.1 Laag 1: awareness

Internetgebruikers kunnen informatie verschaffen aan websites op twee manieren: vrijwillig en niet-vrijwillig. In de eerste vorm, de vrijwillige vorm, vullen internetgebruikers formulieren in of versturen ze email naar exploitanten van website met een vraag of een opmerking. Bij de tweede manier worden gegevens verzameld door deze websites, zonder dat de gebruikers hier op de hoogte van zijn. Het privacyrisico is dan ook dat gegevens ongewild en onwetend worden verstuurd. Veel gebruikers zijn niet op de hoogte van de gevaren. Veel gebruikers weten bijvoorbeeld niet wat een ‘cookie‘ inhoudt. Over het algemeen zijn internetgebruikers er niet van bewust dat met elke muisklik hun eigen ‘persoonlijk‘ profiel wordt bijgewerkt.

3.5.2 Laag 2: control

Door gebruik te maken van speciale software is het mogelijk om automatisch de meest bekende methoden van profiling de baas te zijn. Door speciale plugins in



Figuur 4: Lagen van privacy-awareness, bron: Ishitani [4]

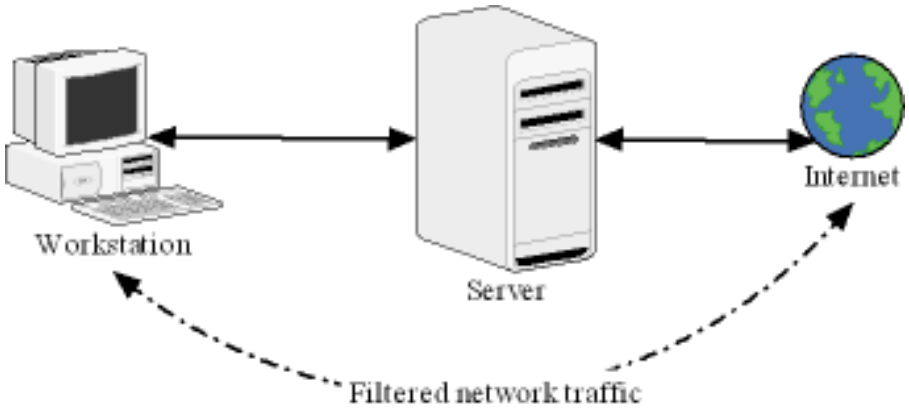
de browser te implementeren worden gevaarlijke stukken code die persoonlijke gegevens verzamelen, geblokkeerd. Grote browsers zoals Microsoft Internet Explorer en Firefox bevatten mogelijkheden om cookies tegen te gaan. Helaas weten erg veel gebruikers niet van het bestaan van deze mogelijkheden en is het uitzetten van cookies al helemaal een onbegonnen zaak doordat ze zichzelf dan echt moeten verdiepen in de applicatie. Ook het gebruik van een anonymizer (zie hoofdstuk 3.3) valt onder deze laag.

3.5.3 Laag 3: Privacy-enhancing tools

Deze laag lijkt veel op de vorige control-laag. Ook hier wordt er gebruik gemaakt van speciale software om profiling tegen te gaan. Het verschil zit hem echter in de locatie van deze speciale software. In laag twee bevindt de software zich op de computer van de gebruiker, de client, zelf. Dit wordt dan ook client-side software genoemd. In laag drie wordt de software opgeslagen en gebruikt op de server. Zie figuur vijf voor een grafische weergave. Het grote voordeel hiervan is dat de eindgebruiker niet meer verantwoordelijk is voor het correct functioneren van deze anti-profiling software en hiermee dus altijd beschermd is tegen aanvallen van buitenaf. De server (zie figuur vijf) regelt al het verkeer tussen het internet en de gebruiker.

3.5.4 Laag 4: Privacy-policies

Het idee achter deze laag is om gebruikers ervan bewust te maken wat het beleid van de desbetreffende website is. Nadat dit duidelijk is, zal de gebruiker een bewuste keuze moeten maken in hoeverre het profilen wordt toegelaten. Een grote



Figuur 5: Client-server-model, bron: R.H.W. Lommers

instantie die hiermee bezig is, is het World Wide Web Consortiums Platform for Privacy Preference Project [P3P, http://www.w3.org/p3p]. Het biedt websites de mogelijkheid om zichzelf te profileren als P3P-complaint; dat wil zeggen dat er door middel van policies exact wordt uitgelegd aan de browser/cliënt wat de bedoeling is van al het dataverkeer dat gegenereerd wordt.

3.5.5 Laag 5: Privacy en trust certification

Deze laag gaat een stap verder dan laag vier; er wordt namelijk vanuit gegaan dat elke website regelmatig haar policies bijstelt en deze up-to-date houdt. Door regelmatig te controleren of de policy nageleefd wordt en dit te beoordelen met cijfers worden websites in categorien gedeeld. Zo worden er cijfers gegeven voor de volgende onderdelen van een website:

- geen geldig verkregen toegang tot cliëntcomputer
- ongeldig verzamelen van gebruikersinformatie
- ongeldig monitoren van gebruikers; d.w.z. het bijhouden van klikgedrag
- ongeldige opslag op cliëntcomputer

Tegenwoordig (anno december 2006) voldoen nog maar weinig websites aan de strenge eisen van P3P. Dit worden er echter wel steeds meer. Helaas is het, net zoals bij de andere lagen, ook hier het geval dat internetgebruikers geen kennis hebben van de software en de mogelijkheden om profilen tegen te gaan.

3.5.6 Laag 6: Privacy-protection laws

In veel landen hebben regeringen gedebatteerd over de mogelijkheden en misbruiken van profilen. Ook zijn er speciale wetgevingen geïntroduceerd die het misbruik van persoonlijke informatie tegen moeten gaan. De laatste laag, laag

zes, gebruikt juridische aspecten om het niveau van profilen vast te leggen. Een probleem is echter wel dat het internet vrij is. Het internet bestaat letterlijk uit miljoenen met elkaar verbonden computers (interconnected), van overal ter wereld. Dit houdt in dat er niet één centrale regering is die regels verzint. Wat in het ene land absoluut niet mag is in het andere land, met een andere regering, de normaalste zaak van de wereld. Een actueel punt is de kwestie over het wel of niet verplichting van providers om het internetgebruik van hun klanten op te slaan.

4 Methode

Er zijn verschillende doelgroepen die dagelijks te maken hebben met profilen. Studenten, arbeiders en hoog opgeleide mensen maken erg veel gebruik van de diensten van het internet. Op welke manier beïnvloedt profiling het gebruik van deze doelgroepen? Dit wil ik gaan onderzoeken voor een bepaalde doelgroep, namelijk de studenten Informatica en Informatiekunde van de Radboud Universiteit, te Nijmegen.

Om de aandacht van de respondenten te krijgen ga ik ze twee verschillende scenario's voorleggen. Het eerste voorbeeld heeft bijna elke internetgebruiker dagelijks mee te maken, namelijk zoekmachines. Het tweede voorbeeld komt minder vaak voor en kan gezien worden als een meer geavanceerdere vorm van profiling.

4.1 Scenario's

Hieronder volgen de verschillende scenarios die ik aan de respondenten voorleg. Pas na het lezen hiervan begint het invullen van de enquête.

4.1.1 De zoekmachine Google

De zoekmachine Google is een dienst van het bedrijf Google, Inc. [http://www.google.com] Met deze dienst is het mogelijk om documenten op het World Wide Web te doorzoeken naar voor jou relevante artikelen. Als jij als gebruiker van Google zoekwoorden invoert, dan is het voor Google mogelijk om deze te groeperen en te koppelen aan een voor jou uniek adres (het IP-adres). Op deze manier verkrijgt Google allerlei informatie die betrekking heeft op jou als gebruiker. Het resultaat hiervan is dat Google erg veel over haar gebruikers te weten komt. Voorbeelden zijn NAW gegevens, interessegebieden, seksuele geaardheid, koopgedrag of medische achtergronden. Met deze gegevens kan Google erg veel bereiken. Zo is het mogelijk om persoonlijke reclame aan te bieden. Dit houdt in dat de reclame die op websites verschijnt op jou persoonlijke toegepast is; de interessegebieden staan immers in jouw persoonlijke profiel. Daarnaast is het mogelijk dat al deze persoonlijke profielen verkocht worden aan derden.

4.1.2 Discriminatie qua inkomen of demografische ligging

Het tweede voorbeeld gaat een stap verder. De door profiling verkregen gegevens worden nu gebruikt om met behulp van decision trees de prijs voor de bezoeker te bepalen. Decision trees kunnen worden omschreven als algoritmes die aan de hand van gegevens bepaalde keuzes maken. Als de bezoeker woont in een gebied met postcode ABCD, dan gaat er een vaste korting van vijf procent van de prijs af. Op deze manier is het mogelijk om bepaalde klanten te

trekken. Stel dat vastgesteld is dat een persoon die een website bezoekt afkomstig is uit het financieel rijke Wassenaar; door middel van descision trees wordt er vervolgens bepaald welke aanbiedingen er in de banners komen te staan. De aanbiedingen liggen in een hogere categorie dan de gemiddelde bezoekers van dezelfde website. Deze vorm van profiling hoeft niet per definitie negatief te zijn, maar kan voor de eindgebruiker toch als negatief overkomen. Dit omdat verschillend mensen ongelijk behandeld worden. Het meest ideale zou zijn om deze optie aan of uit te kunnen schakelen.

5 Enquete

5.1 Informatie

De onderstaande enquête wordt voorgelegd aan studenten van de Radboud Universiteit Nijmegen. De studenten studeren aan de faculteit NIII een studie Informatica of Informatiekunde. Met deze enquête wil ik een antwoord gaan vinden op de hoofdvraag. Deze luidt als volgt:

> *Welk beeld hebben studenten informatica en informatiekunde van de Radboud Universiteit Nijmegen van de mogelijkheden en onmogelijkheden van profilen?*

De deelvragen worden als volgt omschreven.

- Wat is profiling precies?
- Welke soorten van profiling zijn bekend en wat zijn de verschillen?
- Welk soort profiling wordt het meest toegepast?
- Welke doelgroepen komen in aanmerking voor profiling en welke doelgroep is het meest geliefd?

Met behulp van de survey tool 'PHP Surveyor' ga ik een online enquête samenstellen waarmee de respondenten snel en gemakkelijk de vragen kunnen beantwoorden. Deze applicatie genereert een vragenlijst in HTML-code zodat de respondenten op een eenvoudige en bekende wijze de enquête in kunnen vullen.

5.2 Vragenlijst

De volgende lijst met vragen heb ik voorgelegd aan de studenten Informatica en Informatiekunde.

Vraag	**Antwoord mogelijkheden**
01 Wat is uw geslacht?	( ) man ( ) vrouw
02 Wat is uw leeftijd?	( ) 15 t/m 20 jaar ( ) 20 t/m 25 jaar ( ) 25 t/m 30 jaar ( ) 30 t/m 35 jaar ( ) ouder dan 35 jaar
03 Wat is uw hoogst voltooide opleiding?	( ) lager beroeps of algemeen onderwijs ( ) middelbaar beroeps of algemeen onderwijs ( ) hoger beroeps of algemeen onderwijs ( ) wetenschappelijk onderwijs
04 Welke studie volgt u op dit moment?	( ) Informatica ( ) Informatiekunde
05 Waar heeft u de beschikking over internet?	( ) thuis en op werk/school ( ) alleen thuis ( ) alleen op werk/school ( ) Ergens anders ( ) geen
06 Hoeveel dagen per week internet u gemiddeld?	( ) 1 dag ( ) 2 dagen ( ) 3 dagen ( ) 4 dagen ( ) 5 dagen ( ) 6 dagen ( ) 7 dagen

Vraag	**Antwoord mogelijkheden**
07 Geef aan hoeveel uur u per week besteedt aan de volgende toepassingen	* email .... * zoeken naar informatie .... * zo maar wat rondkijken .... * online gaming .... * chatten .... * kopen via internet .... * gokken .... * downloaden .... * nieuwsgroepen .... * erotiek .... * dating ....
08 Kende u het begrip ‘profiling‘?	( ) ja ( ) nee, ga door naar vraag 10
09 Waar heeft u gehoord over het begrip ‘profiling‘?	( ) internet ( ) kranten ( ) televisie ( ) familie/vrienden ( ) collega ( ) anders, namelijk ....
10 Wat houdt volgens u profiling in?	( ) bijhouden van klantinformatie ( ) koppelen van informatie ( ) weet niet ( ) anders, namelijk ....
11 Maakt u zich wel eens zorgen over uw privacy als u gebruik maakt van internetdiensten?	( ) ja ( ) nee, ga door naar vraag 13

Vraag	**Antwoord mogelijkheden**
12 Bij welke diensten maakt u zich zorgen over uw privacy?	( ) email .... ( ) zoeken naar informatie .... ( ) zo maar wat rondkijken .... ( ) online gaming .... ( ) chatten .... ( ) kopen via internet .... ( ) gokken .... ( ) downloaden .... ( ) nieuwsgroepen .... ( ) erotiek .... ( ) dating ....
13 Leest u de privacy-policy's van websites wel eens?	( ) nee, nooit ( ) ja, soms ( ) ja, regelmatig ( ) ja, altijd
14 Zou u een website minder snel raadplegen als de privac-policy niet bekend is?	( ) ja ( ) nee
15 Heeft u wel eens gehoord van p3p?	( ) ja ( ) nee, ga door naar vraag 17
16 Staat uw browser op de standaard security-waarden?	( ) ja ( ) nee
17 Maakt u wel eens gebruik van security software?	( ) ja ( ) nee, ga door naar vraag 19
18 Van welke security software maakt u gebruik?	( ) antivirus applicatie ( ) spam-filter ( ) anti-spyware applicatie ( ) firewall

Vraag	**Antwoord mogelijkheden**
19 In hoeverre schat u uw kennis over de Nederlandse wet op internetprivacy?	( ) eg laag ( ) laag ( ) gemiddeld ( ) hoog ( ) erg hoog
20 Bedankt voor het invullen. Heeft u nog vragen over dit onderzoek?	( ) ja, ga door naar vraag 21 ( ) nee, einde survey
21 Stel hier uw vraag. Vergeet niet uw emailadres te noteren, i.v.m. de reactie op uw vraag.	* ....

6 Interpretatie

6.1 Responsie

Ik heb de elektronische enquête op een speciaal daarvoor ingerichte website[5] geplaatst zodat alle respondenten op een eenvoudige manier de vragenlijst konden invullen. Het softwarepakket "PHPSurveyor"[6] maakt het mogelijk om op eenvoudige wijze een enquête samen te stellen, die vervolgens door middel van een web-interface benaderd kan worden. De enquête heeft in totaal 2,5 weken online gestaan. Tijdens deze weken hebben 46 personen de enquête ingevuld. Het aantal respondenten is in mijn onderzoek dus niet groot genoeg om een diepgaande kwantitatieve betekenis te vormen over de relatie tussen enerzijds de studenten Informatica- en Informatiekunde en anderzijds het fenomeen 'profiling'. Wel kan het onderzoek gebruikt worden als basis om in de toekomst verder onderzoek te verrichten naar het profiling onder een bepaalde groep mensen.

Het interpreteren van de onderzoeksresultaten ga ik onderverdelen in drie fasen, namelijk:

1. samenvatting resultaten
2. correlaties
3. koppeling met privacy-awareness

6.1.1 Samenvatting resultaten

- Van alle 46 respondenten waren 45 personen van het mannelijke geslacht en één persoon van het vrouwelijke geslacht.
- Het grootste deel, namelijk 32 (69,57%) personen, viel in de leeftijdscategorie 20 tot 25 jaar.
- Als vooropleiding had 50% een HAVO/HBO opleiding achter de rug, 26% een VWO/WO en 19% was afkomstig van de MAVO/MBO.
- Het onderzoek was gericht op studenten van de faculteit NIII[7] van de opleidingen Informatica en Informatiekunde. Dertien personen (28,26%) studeerden Informatica en 33 personen (71,74%) Informatiekunde.
- Op de vraag waar de respondenten de beschikking hadden over een internetaansluiting antwoordden 100% dat zij een aansluiting hadden op hun thuislocatie en op de universiteit. 50% beschikt op hun (parttime)werk over een verbinding met het internet en geen enkele respondent beschikte over geen enkele verbinding.

[5] http://nijmegen.lommersonline.com/projects/phpsurveyor/index.php?sid=1
[6] http://sourceforge.net/projects/phpsurveyor/
[7] Nijmeegs Instituut voor Informatica en Informatiekunde

- Ruim 80% van de ondervraagden maakt zeven dagen per week gebruik van internet; 17% doet dit zes keer per week en 3% vijf keer.

- Op de vraag of het begrip "profiling"bekend is, antwoordt 60,87% positief, de overige 39,13% had nog nooit gehoord van profiling. Van deze 60% die wel hadden gehoord van profiling, heeft meer dan de helft deze kennis vergaard door middel van het internet; de rest via de televisie of familie, vrienden en kenissen.

- De meningen over wat nu precies profiling inhoudt verschillen nogal: 45% denkt dat het het bijhouden van klantinformatie is en 44% is van mening dat het te maken heeft met het koppelen van informatie. Slechts 4% weet niet wat profiling inhoudt.

- Op de vraag of men zich wel eens zorgen maakt over privacy tijdens het gebruik van internetdiensten antwoordt 73,91% postitief en 26,09% negatief. Van de respondenten die zich zorgen maken zijn online shopping (58%), email (27%), zoeken naar informatie (36%), downloaden (30%) en chatten (28%) de grootste boosdoeners.

- Privacy-policy's worden door 43% niet gelezen en 47% leest ze soms. Ook maakt het 70% van de respondenten niet uit of een website wel een privacy-policy bevat en heeft 93% nog nooit van p3p[8] gehoord.

- Van de mensen die p3p wel kennen heeft 75% de webbrowser zo ingesteld dat hiervan gebruik wordt gemaakt.

- Er wordt veel gebruik gemaakt van security software. Maar liefst 95,65% heeft een of meerdere softwarepakketten zoals een firewall, spamfilter, antivirus of anti-spywarepakket geïnstalleerd staan.

- De eigen kennis van de Nederlandse wet op internetprivacy wordt door 50% van de de respondenten geschat op gemiddeld.

6.1.2 Correlaties

Alle respondenten hebben laten weten dat ze zowel op de thuisplek als op de universiteit toegang hebben tot internet. Ook wordt er ontzettend veel gebruik van gemaakt. Van studenten Informatica en Informatiekunde valt dit ook wel te verwachten. Maar liefst 97 procent maakt minimaal zes dagen per week gebruik van internet en de diensten die het internet biedt. Wat wel opvalt is dat 40 procent nog nooit had gehoord van profiling terwijl de vraag 'wat houdt volgens u profiling in' door 45 procent wordt beantwoord met 'bijhouden van klanten-informatie' en 43 procent denkt dat het 'koppelen van informatie' is. Ook maakt ruim 73 procent zich wel eens zorgen over over hun privacy tijdens het surfen. Dit verklaart het grote gebruik in softwarepakketten die ervoor dienen ter bescherming van de privacy zoals firewalls, antivirus- en antispyware paketten. Wel vreemd is het dat bijna geen enkele respondent gehoord heeft

[8]Platform for Privacy Preference Project, http://www.w3.org/p3p

van het Platform van Privacy Preference Project (p3p). De mensen die het wel kennen, hebben dan ook laten weten dat zij hun browser zo hebben ingesteld dat er optimaal gebruik wordt gemaakt van de p3p-eigenschappen.

6.1.3 Koppeling met privacy-awareness

Door het plaatsen van de respondenten in één of meerdere privacy-lagen (zie hoofdstuk 3.4) probeer ik een beeld te vormen van het bewust zijn van profiling onder de studenten Informatica en Informatiekunde van de Radboud Universiteit, te Nijmegen. Ik ga per laag bespreken hoe de verhouding is met de respondenten.

1. Laag 1: Awareness. Ruim 70 procent van de respondenten maakt zich wel eens zorgen over hun privacy als zij over het internet surfen. Ook het feit dat 95 procent een security-applicatie geïnstalleerd hebben op hun computer maakt het mogelijk om ze sowieso te plaatsen in de eerste laag: awareness. De respondenten zijn zich ervan bewust dat profiling wordt toegepast en zij kennen de bijkomende gevaren.

2. Laag 2: Control. Op twee respondenten na gebruikt iedereen speciale software voor het garanderen van de security. Het softwarematig tegengaan van profilen staat centraal in de tweede laag. Dit maakt het mogelijk om ze in de category van de tweede laag (control) te plaatsen. Wel is het zo dat de doelgroep studenten Informatica of Informatiekunde zijn. De kans is hierdoor groot dat zij sowieso bekend zijn met softwarematige oplossingen als firewalls en antispam-software. Ook het juist kunnen configureren en instellen hiervan moet geen problemen opleveren bij deze doelgroep.

3. Laag 3: Privacy-enhancing Tools. Deze laag lijkt veel op de vorige laag. Het grote verschil zit hem in de locatie van de tools die ervoor zorgen dat de privacy gegarandeerd blijft. Bij de tweede laag is dit de computer van de gebruiker zelf en bij deze laag gebeurt dit op een speciale server. Zie hoofdstuk 3.4.3 voor meer informatie over het client-server-model. Alle respondenten kunnen in laag drie geplaatst worden met als reden dat ze email-functionaliteiten van de Radboud Universiteit gebruiken. Deze mailserver controleert alle binnenkomende en uitgaande mail automatisch op virussen waardoor cliënt-side controle veel minder noodzakelijk wordt.

4. Laag 4: Privacy Policies. P3P Is niet bekend bij het overgrote deel van de respondenten. Slechts zes procent van de ondervraagden heeft wel eens gehoord van p3p en heeft hun browser zo ingesteld dat deze er gebruik van maakt. Ook worden de policy's van websites zelden gelezen. Dit is dan ook de reden dat er geen respondenten zijn die passen in deze laag van het profiling-awareness-model.

5. Laag 5: Privacy and Trust Certification. Doordat de policies weinig tot niet worden gelezen kunnen de studenten niet worden geplaatst in de

vijfde laag van het model. Een verklaring hiervan kan zijn dat het p3p-model nog niet door veel websites gebruikt wordt. Het is mogelijk dat de studenten er meer gebruik van gaan maken als het breder ondersteund en geaccepteerd wordt. Het zijn nu vaak de 'kleine lettertjes' die zo snel als mogelijk weg worden geklikt en respondenten ervaren ze eerder als lastig en negatief dan ter voorkoming van profiling.

6. Laag 6: Privacy-Protection Laws. In Nederland verschijnt de kwestie 'Online Privacy' steeds vaker in de media. Tijdens dit onderzoek was er een documentaire over Google op het actualiteitenprogramma Zembla[9]. Ook zijn de plannen om alle serviceproviders in Nederland te verplichten tot het opslaan van internetgegevens realiteit. Dit omdat er op Europese schaal gewerkt wordt aan richtlijnen waardoor alle bij de EU-aangesloten landen worden verplicht tot het bijhouden van internet- en telefoongegevens. Deze opgeslagen gegevens kunnen een rol spelen bij onder andere terreurbestrijding. Of dit wetsvoorstel erdoor komt valt nog over te discussiëren. Voor kleine ondernemingen als providers en/of telecombedrijven wordt het een te kostbare zaak om alle gegevens vast te leggen. De reden echter dat de respondenten wél in laag zes te plaatsen zijn, is dat de overheid zich wel degelijk bezighoudt met de kwestie Online Privacy.

6.2 Conclusie

Mijn verwachtingen over het bewustzijn van studenten Informatica en Informatiekunde van het fenomeen profiling waren redelijk gelijk aan de resultaten. Mijn gedachten dat juist deze groep studenten kennis hebben van de mogelijkheden van internet, en dan met name voor profiling met commerciële doeleinden, komt overeen met de resultaten van dit onderzoek.

Na het afnemen van de interviews bleek dat het begrip profiling wel degelijk bekend is bij 60 procent van de studenten. De gemiddelde ondervraagde student kan, aan de hand van de door hem/haar toegepaste middelen, geplaatst worden in de lagen 1, 2, 3 en 6 van de privacy-awareness schaal. Hieruit volgt dan ook dat de studenten zich wel degelijk bewust zijn van profiling, omdat ze zich wél zorgen maken over hun privacy en hier ook actief wat aan doen. Het gebruik van security-software zoals firewalls, antivirus- en spamfilters is hoog omdat de studenten bang zijn voor de gevolgen van profiling.

Aan de andere kant zijn er ook resultaten te bespeuren die tegenstrijdig zijn. Zo maakt ruim driekwart van de ondervraagden zich wel eens zorgen over privacy, terwijl bijna niemand de privacy-policyies van websites leest. De grootste issue is de moeite die het kost om veilig te kunnen internetten. Ze zijn wel bewust van de gevaren, maar het lezen van een A4-tekst kost teveel

[9]Zembla is een actueel documentaire programma van de VARA en NPS (http://redir.vara.nl/zembla/)

moeite. Het kost moeite om de policy's te lezen, terwijl het gebruik van security-software geen enkele moeite kost. Alleen de installatie ervan vergt enige minuten tijd.

Mijn conclusie is dat de respondenten de gevaren van profiling niet groot genoeg vinden/kennen om deze maximaal te bestrijden. Simplele handelingen als de installatie van security software worden nog wel verricht, maar het lezen van privacy-policies en/of het instellen van uitgebreide browserinstellingen voor p3p is te veel werk. Wellicht als p3p meer bekendheid krijgt, wordt het mogelijk om de browser-software hierop aan te passen zodat nog meer internetgebruikers in de lagen vier en vijf van het model geplaatst kunnen worden. Ook denk ik dat veel mensen niet in de lagen vier en vijf zitten omdat ze de gevaren van profiling onderschatten. Het grootste deel denkt dat profiling inhoudt dat gegevens worden bijgehouden en is zich er niet van bewust dat dit veel verder kan gaan. Als het begrip wat meer aandacht krijgt, wat overigens de laatste tijd steeds vaker het geval is, dan verwacht ik dat meer mensen zich gaan bestrijden tegen de gevaren van profiling.

Tevens schuilt er zich een probleem in de mogelijkheden van profiling. Het wordt namelijk niet enkel gebuikt voor negatieve doeleinden. Ook het opsporen van fraude, corruptie, medische diagnose en wetenschappelijke ontdekkingen zijn mogelijkheden van profiling/datamining. Als gebruikers actief deze techniek tegengaan verdwijnen ook deze positieve resultaten. Zoals eerder gezegt is het dan ook zaak dat er meer aandacht wordt besteedt aan de gevolgen van profiling. Mensen moeten op de hoogte worden gebracht van wat er allemaal mogelijk is en hoe ze dit tegen kunnen gaan. Een goed vervolgonderzoek op deze bachelorscriptie zou kunnen zijn om uit te zoeken hoe ver mensen willen gaan tot het openstellen van privacygevoelige informatie en waarom ze dat willen.

7 Slotwoord

De techniek staat voor niets wordt er vaak gezegd. De mogelijkheden van profiling zijn in een paar jaar tijd enorm gegroeid. Ik denk dat het belangrijk is dat internetgebruikers de gevaren van dit profilen inzien en hier rekening mee houden. Met deze bachelorscriptie heb ik getracht inzicht te krijgen in het begrip profiling. Vervolgens heb ik onderzocht hoe studenten Informatica en Informatiekunde van de Radboud Universiteit in Nijmegen bewust zijn van de gebruikte profiling-technieken. Via deze weg wil ik alle mensen bedanken die mee hebben gewerkt aan het tot stand komen van deze scriptie. In het bijzonder wil ik Dhr. Luca Consoli bedanken voor zijn tijd en inzet tijdens het begeleiden van mijn scriptie-periode.

Rogier Lommers, april 2006
rogier@lommersonline.com

Referenties

[1] S. Spiekermann, I. Dickinson, O. Gunther: *User Agents in E-commerce Environments: Industry vs. Consumer Perspectives on Data Exchange*, http://portal.isiknowledge.com/portal.cgi, bekeken op 12/08/2005.

[2] Y. Elovici, C. Glezer, B. Shapira: *Enhancing customer privacy while searching for products and services on the World Wide Web*, http://www.emeraldinsight.com/Insight, bekeken op 02/11/2005.

[3] Google Inc. :*Google Privacy Policy*, http://www.google.com/, bekeken op 04/11/2005.

[4] L. Ishitani, V. Almeida, W. Meira: *Masks: Bringing anonymiti and personalization together*, http://ieeexplore.ieee.org/iel5/8013/27102/01203218.pdf, bekeken op 15/11/2005.

[5] B. Schneier: *Secrets and Lies: Digital Security in a Networked World*, ISBN: 0-471-25311-1.

[6] C. Arndt: *The loss of privacy and Identity*, http://www.comdig.org/article.php?id_article=22829, Ideal Innovations Inc., bekeken op 15/11/2005.

[7] E. Taylor_Powell: *Questionnaire Design: Asking question with a purpose.*

[8] W. Dijkstra en J. H. Smit: *Onderzoek met vragenlijsten*, VY Uitgererij Amsterdam, 1999.

[9] Future of Identity in the Information Society: *Descriptive analysis and inventory of profiling practices*, http://www.fidis.net, bekeken op 15/12/2005.

[10] S. Fisher, D. Thomas: *Privacy and security at risk in the global information society*, 2000.

[11] D. Sellars: *An Introduction to Steganography*, University of Capetown, 1999.

[12] V. Senicar, B. Jerman, T. Klobucar: *Privacy_enhancing technologies: approaches and developments*, Computer Standards and Interfaces, Vol. 25, pages 147-158.

[13] M. Driscoll, C. Roberts, E.Lyons, J. Jain, J. Nuckols: *Secure Online Payment Systems*, http://mba.vanderbilt.edu/securepayments/frames.html, bekeken op 19/12/2005.

[14] J. Claessens, B. Preneel, J. Vandewalle: *Solutions for anonymous communication on the internet*, Proceedings of the 1999 IEEE

International Carnahan Conference on Security Technology, pages 298-303.

[15] Michael J. Shaw, Chandrasekar Subramaniam, Gek Woo Tan, Michael E. Welge: *Knowledge Management and Data Mining for Marketing*, Elsevier Science B.V., 2001.

[16] Wikipedia, de Vrije Encyclopedie :*Datamining*, http://nl.wikipedia.org/wiki/Datamining, bekeken op 09/06/2006.

Lijst van figuren

A Bijlagen

A.1 Respondentie interviews