

User Profile Matching

Yannick Van de Weerd

promotor :
Prof. dr. Karin CONINX

Abstract

Gebruikersgerichte informatie aanbieden aan geïnteresseerde gebruikers kan nuttig zijn in vele omstandigheden. Hierbij denken we aan reclame die enkel aan mogelijke kopers gepresenteerd wordt zonder gebruikers nodeloos lastig te vallen met ongewenste reclameboodschappen. Een andere toepassing is om personen aan elkaar voor te stellen door persoonlijke informatie in rekening te brengen bij het zoeken naar interessante personen. Personen kunnen interessant zijn in verschillende contexten. Deze context kan een zakenomgeving zijn, maar ook een vrijetijdsomgeving waarbij een interessant persoon, iemand is die eenzelfde sport beoefent. Deze laatste toepassing hebben we in deze thesis nader bestudeerd door een literatuurstudie en een implementatie die de beschreven technieken aan de praktijk toetst.

Woord vooraf

Een thesis maken is meer dan een taak waarbij de universiteit een onderwerp opgeeft en waarover de student een 70-tal pagina's moet schrijven. Het maken van een thesis vergt veel creativiteit gecombineerd met enthousiasme en tijd. Daarom is het van groot belang dat het onderwerp de student aanspreekt zodat algemene interesse en thesisonderwerp elkaar ontmoeten. In deze thesis heeft deze ontmoeting plaatsgevonden hetgeen een positieve invloed heeft gehad op het uiteindelijke resultaat.

Voor een geslaagde thesis zijn meer ingrediënten nodig dan enkel een boeiend onderwerp. Vrienden, collega's, familie en een goede begeleiding vormen het hoofdingrediënt van een thesis. Daarom wil ik bij deze graag iedereen bedanken die heeft meegeholpen bij de realisatie van deze thesis.

Tim Clerckx en Yves Vandriessche bedank ik voor de uitstekende begeleiding. Zowel Tim als Yves hebben geholpen bij het in goede banen leiden van de thesis. Prof. dr. Karin Coninx wil ik ook bedanken voor het mogelijk maken van deze thesis. Verder wil ik ook vrienden en collega's bedanken zoals Ruben Thys en Tom Cuypers voor de brainstorming en latex hulp. Voor het nalezen van deze thesis wens ik graag Brigitte Nolens te bedanken.

Inhoudsopgave

1	Inleiding	1
1.1	Domein	1
1.2	Probleemstelling	2
1.3	Doel	3
1.4	Slaagkansen in de praktijk	4
1.5	Opdeling thesis	4
2	Gerelateerd werk	6
2.1	Implicit Human-Human Interaction	7
2.1.1	Digital Aura	7
2.1.2	Gegevens uitwisseling	9
2.2	Social Proximity Sensing	9
2.3	Sociaal Perspectief	11
2.4	Gebruikersprofiel	12
2.4.1	Richtlijnen	12
2.4.2	UserML	13
2.4.3	IHHI	13
2.4.4	FOAF-profiel	15
2.4.5	Andere mogelijkheden	16
2.5	Intelligente Profielen	16
2.6	Privacy and Trust	18
2.7	Social Network Sites	20
2.8	Conclusie	22

3	User Profile	23
3.1	Situering	24
3.2	Profile Designer	26
3.3	Soorten variabelen	26
3.4	User Profile Richtlijnen	27
3.4.1	Gebruiksgemak	28
3.4.2	Dubbelzinnige informatie	30
3.4.3	Standaardwaarden	30
3.4.4	Verplichte velden	31
3.4.5	Wijzigingen toelaten	31
3.4.6	Exacte data en aanliggende profielitems	32
3.4.7	Ordering	33
3.4.8	Belangrijke elementen	34
3.5	Context Awareness	35
3.6	Conclusie	37
4	User Profile Matching	38
4.1	User Profile Matching Structuur	39
4.2	Analyse en opsplitsing van het probleem	39
4.3	Basic Matching	42
4.4	Foaf Matching	45
4.5	Trusted Matching	48
4.6	Conclusie	51
5	Implementatie	54
5.1	Website	55
5.1.1	Inloggen	56
5.1.2	Aanpassen van profielinformatie	56
5.1.3	Aanpassen van zoekprofiel	58
5.1.4	Aanpassen van profielelementen door profile editor	58
5.1.5	Zoeken naar gebruikers	61
5.1.6	Beheren van contacten	63

5.1.7	Profielen van andere gebruikers bekijken	65
5.1.8	Exporteren contacten naar PDA applicatie	65
5.1.9	Conclusie Website	65
5.2	Serverapplicatie	65
5.2.1	Network Users & Network Graph	66
5.2.2	Selected User	68
5.2.3	Component Weights	68
5.2.4	Components Info	69
5.2.5	Calculation	69
5.2.6	Synchronizatie met website	70
5.3	Synchronisatie Server	70
6	Conclusies	72
7	Future Work	76
7.1	Intelligent zoekprofiel	77
7.2	Digital Aura op gsm	77
7.3	User Tests	81

Lijst van figuren

2.1	Overlapping tussen twee aura's [21].	7
2.2	Scenario waarbij twee zakenmensen elkaar ontmoeten op een receptie en een scenario waarin een werknemer op zoek is naar een tegenspeler.	8
2.3	Een taxonomie van een gebruikersmodel om mee te starten. .	14
2.4	Voorbeeld van een gebruiksprofiel dat door IHHI gebruikt wordt.	15
2.5	Voorbeeld RDF bestand.	17
2.6	Voorbeeld van een sociaalnetwerkstructuur.	20
3.1	Voorbeeld van een profielement.	28
3.2	Voorbeeld van een profiel waarbij profielementen gegroepeerd kunnen worden en waar niet-ingevulde profielementen niet worden getoond.	29
3.3	Grafieken die aangeven hoezeer aanliggende profielitems meetellen in de berekening.	32
3.4	Ordening van de profielgroepen.	34
3.5	Het geven van een score aan een profielement door een star-rating mechanisme.	35
4.1	Eenvoudige voorstelling van een chain waarin verschillende componenten doorlopen worden om een uitkomst te bekomen.	40
4.2	Opsplitsing van berekening chain over verschillende CPU's. .	40

4.3	Een profielement met bijbehorende profielitems waaruit gekozen kan worden.	43
4.4	Een profielement met bijbehorende waarderingscore.	43
4.5	Grafische representatie van de foaf-afstand tussen personen. .	45
4.6	Illustratie van de trust berekening tussen twee personen die elkaar rechtstreeks kennen.	48
4.7	Illustratie van de trust berekening waarbij gebruik gemaakt wordt van een tussenpersoon.	49
4.8	Resultaat van een trust-score berekening waarbij gebruik gemaakt wordt van een tussenpersoon.	50
4.9	Illustratie van de trust berekening waarbij een gebruiker een andere gebruiker een nul-score geeft.	50
4.10	Illustratie van de trust berekening waarbij verschillende paden mogelijk zijn.	51
5.1	Bewerk bestaande groepen en items en voeg groepen en items toe aan het profiel.	57
5.2	Profielementen zijn opgedeeld in de categorieën Identity, Work, Love en FreeTime.	59
5.3	Aanpassen van profielementen door een Profile Designer. . .	60
5.4	Grafiek editor om de relatie tussen de score en de profielitems aan te geven.	61
5.5	Webpagina om naar andere gebruikers in het netwerk te zoeken.	62
5.6	Beheren van de contacten en toekennen van een trust score. .	64
5.7	Overzicht van de serverapplicatie.	66
5.8	Overzicht van alle netwerkgebruikers.	67
5.9	Graaf van alle netwerkgebruikers.	67
5.10	Profiel van de geselecteerde gebruiker.	68
5.11	Illustratie van de gewichten die aan de componenten gekoppeld zijn.	68
5.12	Overzicht van de berekeningen benodigd voor een component.	69

5.13	Uitkomst van iedere component, vermenigvuldigd met een gewicht om tot een eindresultaat te komen.	69
5.14	Informatie over de aanvragen via de website.	70
5.15	Synchronisatie-server voor het synchroniseren van de contacten tussen de site en de PDA applicatie.	71
5.16	Overzicht van de synchronisatie werking.	71
7.1	Contactenlijst xml structuur.	80

Hoofdstuk 1

Inleiding

Inhoudsopgave

1.1	Domein	1
1.2	Probleemstelling	2
1.3	Doel	3
1.4	Slaagkansen in de praktijk	4
1.5	Opdeling thesis	4

1.1 Domein

Het domein van deze thesis kunnen we best illustreren aan de hand van een voorbeeld scenario. Stel een werknemer die een werkgerelateerd probleem in een specifiek domein heeft. Stel dat deze werknemer deze informatie kan verspreiden. Op deze wijze kunnen andere werknemers die bekend zijn met het gevraagde domein deze vraag opmerken en eventueel beantwoorden. Door deze automatisatie kan een bedrijf veel tijd uitsparen.

Er zijn talloze andere scenario's denkbaar waarbij personen op zoek zijn naar andere personen. Om deze zoektochten te automatiseren moeten we

verschillende onderzoeksdomeinen combineren. Een eerste domein is Ubiquitous Computing. Ubiquitous Computing is een visie waarbij computersystemen overal aanwezig zijn en transparant met elkaar en met de gebruiker samenwerken. Transparantie voor de gebruiker betekent vooral dat het gebruik van deze computersystemen niet opgemerkt wordt.

Een tweede domein is het beschrijven van personen en hetgeen naar wat ze op zoek zijn zodat een computer deze informatie kan verwerken. Het beschrijven van een gebruiker zodat een computer dit begrijpt zal gebeuren in een User Profile. Het zoeken naar gebruikers die voldoen aan een bepaalde beschrijving behandelen we onder User Profile Matching. In deze thesis zal onderzoek gedaan worden naar een User Profile en de bijhorende User Profile Matching. Het nut van profile matching is snel duidelijk als het een gegeven is dat de populariteit van dating sites in de USA met 50 % is gestegen [4]. Deze enorme stijging geeft duidelijk aan dat de computerleek de weg naar de online sociale netwerken, waarin de gebruikers aan elkaar gematcht kunnen worden, gevonden heeft.

Een ander voorbeeld is de enorme toename van het gebruik van social bookmarking systemen zoals del.icious.com, digg.com, Dit geeft aan dat gebruikers informatie willen delen met elkaar om zo kennis te verspreiden en zelf makkelijker gerichte informatie te kunnen verkrijgen.

Bij het bovenstaand scenario van de werknemer valt het automatisch zoeken naar de juiste persoon onder user profile matching, hierbij helpt het ubiquitous computing onderzoek bij het transparant uitvoeren van deze matching. In volgende sectie beschrijven we kort wat de problemen bij een profile matching kunnen zijn.

1.2 Probleemstelling

Een systeem dat op zoek gaat naar interessante gebruikers zoals beschreven in bovenstaand scenario kan enkel goed werken indien gebruikers zichzelf

en hetgeen naar wie ze op zoek zijn in voldoende detail kunnen beschrijven. Om een persoonsbeschrijving, User Profile genoemd, te omschrijven zodat een computer dit begrijpt zullen we behandelen in deze thesis. In volgende hoofdstukken zullen we enkele richtlijnen geven om een goede User Profile op te stellen.

Er moet een onderscheid gemaakt worden tussen de interessante gebruikers en de niet-interessante gebruikers. Om het verschil tussen interessante en niet-interessante gebruikers aan te tonen maken we gebruik van bovenstaand scenario. In bovenstaand scenario is een weknemer die kennis heeft van het gevraagde domein een interessante persoon. Indien deze persoon te werk gesteld is in een concurrerend bedrijf, is deze persoon een niet-interessant persoon. Bovendien is een persoon vaak interessant naargelang de context waarin de gebruiker zich bevindt. Zo kan een mogelijk interessante gebruiker voor bovenstaand scenario op vakantie zijn. In dit geval storen we deze persoon liever niet met werkgerelateerde vragen. Dit is waar de User Profile Matching in werking zal treden zodat de informatie naar de juiste personen gestuurd wordt.

1.3 Doel

Het doel van dit werk is om zowel User Profile richtlijnen als User Profile Matching richtlijnen te geven. Om aan deze eis te voldoen moeten we ons onderzoek opsplitsen in verschillende aspecten die we vervolgens laten samenwerken als één geheel. Zo moet er onderzoek gedaan worden naar een methode om een gebruiker te beschrijven in de vorm van een gebruikersprofiel (hoofdstuk 3), een methode om de gebruikersprofielen uit te wisselen (sectie 2.1.1) en een methode om interessante profielen van de minder interessante profielen te onderscheiden (hoofdstuk 4), User Profile Matching genoemd. Bij het zoeken naar de meest geschikte methode houden we rekening met de uitbreidbaarheid en het universeel zijn van de aangeboden technieken. Daarenboven willen we dat het mogelijk is om het gebruikersprofiel te exporteren naar een formaat dat leesbaar is door het Digital Aura

framework, ontwikkeld aan het EDM [49], [7].

1.4 Slaagkansen in de praktijk

Onderzoek naar nieuwe technieken en mogelijkheden om interessante profielen van de minder interessante profielen te onderscheiden, User Profile Matching genoemd, kan boeiende resultaten opleveren. De resultaten die we in deze thesis behalen kunnen makkelijk in de praktijk toegepast worden, en dat maakt het onderzoek meer concreet. In hoofdstuk 5 over de implementatie wordt aangetoond hoe deze thesis naar de praktijk omgezet kan worden.

In ons systeem worden beschrijvingen van gebruikers (zie hoofdstuk 3) met elkaar vergeleken op een Social Network Site (zie sectie 5.1) om vervolgens de gematchte personen over te zetten naar bijvoorbeeld een PDA-toestel dat de gebruiker een bericht toont indien een interessante gebruiker in de buurt komt.

1.5 Opdeling thesis

Deze thesis is opgedeeld in enkele grote hoofdstukken die ieder een verschillend aspect van User Profile Matching behandelen. In het eerste hoofdstuk onderzoeken we de bestaande implementaties die deels of volledig overeenstemmen met ons onderzoeksdomein. Deze implementaties analyseren we om vervolgens conclusies te maken en eventueel stukken of ideeën te herbruiken in onze uitwerking.

Het tweede hoofdstuk behandelt de structuur en vereisten van een gebruikersprofiel. Opdat een profiel voldoende informatie over de gebruiker kan bevatten is het van belang dat de structuur toelaat om makkelijk uitgebreid en aangepast te worden.

Het daaropvolgend hoofdstuk behandelt de profile matching op zich. Hier worden profielen met elkaar vergeleken en wordt de ketting structuur voorgesteld. Deze matching bestaat uit drie grote delen. Het eerste deel is een profiel matching waarbij profielelementen met elkaar vergeleken worden. Het tweede deel maakt een verschil tussen kenissen en vreemden. De trust component tracht een waarde over de betrouwbaarheid te berekenen. Deze componenten worden gebundeld om vervolgens een score terug te geven.

Hoofdstuk 5 legt kort de implementatie uit en toont aan dat hetgeen besproken is in voorgaande hoofdstukken toepasbaar is in de praktijk. De implementatie bestaat uit een server waar de matching score berekent wordt en een website vanwaar gebruikers een profiel kunnen opstellen en aangeven naar welke personen ze op zoek zijn.

Tot slot bespreken we wat de toekomstige mogelijkheden van deze thesis zijn. Zo geven we aan dat een user test waarschijnlijk interessante resultaten oplevert en dat een digital aura relatief eenvoudig gemaakt kan worden door gebruik te maken van gsm's. Verder geven we aan dat een intelligent zoekprofiel een grote meerwaarde voor ons systeem zou betekenen.

Hoofdstuk 2

Gerelateerd werk

Inhoudsopgave

2.1	Implicit Human-Human Interaction	7
2.1.1	Digital Aura	7
2.1.2	Gegevens uitwisseling	9
2.2	Social Proximity Sensing	9
2.3	Sociaal Perspectief	11
2.4	Gebruikersprofiel	12
2.4.1	Richtlijnen	12
2.4.2	UserML	13
2.4.3	IHHI	13
2.4.4	FOAF-profiel	15
2.4.5	Andere mogelijkheden	16
2.5	Intelligente Profielen	16
2.6	Privacy and Trust	18
2.7	Social Network Sites	20
2.8	Conclusie	22

2.1 Implicit Human-Human Interaction

Computers zijn alomtegenwoordig in onze samenleving. De nood aan computers die met elkaar samenwerken zonder dat de gebruiker (veel) input moet geven wordt steeds groter. In de literatuur wordt dit probleem omschreven als Ubiquitous Computing.

Schmidt [43] definieerde implicit interaction als een nieuw interactiemodel waar interactie gebaseerd is op geluid en locatiegegevens. Clerckx [49] vertrekt van dit model en creëert de “implicit human-human interaction”. Dit model beschrijft hoe personen met elkaar communiceren via impliciete interactie. Dit project maakt gebruik van een PDA-systeem om de gebruiker te berichten over andere gebruikers die in de buurt zijn.

2.1.1 Digital Aura



(a) Overlappend van twee aura's
(b) Voorbeeld van een overlapping van aura's in een reële omgeving.

Figuur 2.1: Overlapping tussen twee aura's [21].

Digital Aura is een mogelijke oplossing om een interactie tot stand te brengen als twee gebruikers zich in elkaars buurt bevinden. Digital Aura maakt gebruik van een metaforische aura die zich rond een gebruiker bevindt. Bij een overlapping van twee aura's, zoals geïllustreerd in figuur 2.1(a), waarbij één van de twee bronnen zich in de intersectie bevindt, ontstaat er interactie tussen beide bronnen. Merk op dat de tekening dus in-



Figuur 2.2: Scenario waarbij twee zakenmensen elkaar ontmoeten op een receptie en een scenario waarin een werknemer op zoek is naar een tegenspeler.

correct is. Om te figuur niet te overladen hebben we de personen niet in de intersectie getekend.

Ferscha [21] heeft als voorbeeldscenario een student die op zoek is naar hulp voor een bepaald studievak. De student maakt deze vraag kenbaar in de toepassing. Bij het kruisen van de lesgever ontvangt de lesgever een bericht dat hij de betreffende student met een probleem gekruist heeft. Desgewenst kan de lesgever reageren op dit bericht en de student contacteren.

Een ander voorbeeldscenario is een werknemer die is tewerkgesteld op een groot bedrijf en graag zou gaan squashen (Figuur 2.2). De werknemer zit met het probleem dat hij geen tegenspeler vindt. Door in de toepassing aan te geven dat hij een tegenspeler voor squash zoekt, zien medewerknemers deze vraag en kunnen medewerkers contact opnemen om concrete afspraken te maken.

In de bedrijfswereld kan een digitalauratoepassing ook zijn nut bewijzen (Figuur 2.2). Door als zakenman aan te geven waarnaar hij op zoek is, wat hij aanbiedt en wie hij is, kunnen andere zakenmensen, die mogelijk

geïnteresseerd zijn, automatisch een bericht ontvangen indien deze zakenman zich in de buurt bevindt. Een extra functionaliteit die deze toepassing kan bieden is het tonen van een gemeenschappelijke interesse, om zo als ijsbreker te dienen die de communicatie vergemakkelijkt.

2.1.2 Gegevens uitwisseling

Indien twee aura's met elkaar intersecteren, worden er gegevens uitgewisseld. De uitwisseling van gegevens dient voldoende flexibiliteit te bieden om aan een voldoende grote diversiteit aan scenario's ondersteuning te bieden. De gegevens die uitgewisseld worden, zijn in een gebruikersprofiel beschreven. Dit profiel wordt uitgebreid beschreven in hoofdstuk 3 over gebruikersprofielen.

Bij een automatische uitwisseling van gegevens moet erop gelet worden dat de gebruiker zijn privacy niet in het gedrang komt. Om aan deze eis te voldoen wordt privacyinformatie mee opgenomen in het profiel. Bij de profile matching kunnen we vervolgens rekening houden met deze privacyinformatie. Hier komen we uitgebreid op terug in sectie 2.6 over Privacy and Trust.

2.2 Social Proximity Sensing

Om de digital aura technologisch te verwezenlijken zijn er tal van technieken ontwikkeld. Normaliter maakt men gebruik van apparatuur die een beperkt bereik heeft zodat de digital aura automatisch gecreëerd wordt door de technologische beperkingen.

In 1998 ontstond de Lovegety [24] in Japan. Deze tool maakt gebruik van een radiofrequentie om gegevens tussen de verschillende Lovegety's uit te wisselen. De uitwisselingsafstand bedraagt 5 m. Het gebruikersprofiel is zeer beperkt. Het profiel bestaat uit het geslacht van de gebruiker en de stemming (vrolijk, verdrietig, ...) van de gebruiker op dat moment. Als

twee personen elkaar tegenkomen met een gelijke stemming en verschillend geslacht geeft de applicatie een notificatie.

Iets later komt Gaydar tot leven in de United States. Dit apparaat vertoont grote gelijkenissen met de Lovegety, maar was gericht op de halebivereniging.

Er is een grote variëteit in de ontwikkelde systemen voor *Social Proximity Sensing*. Sommige systemen zijn reeds commercieel verkrijgbaar en worden positief ontvangen door de gebruikers ervan. De doelstelling van deze systemen zijn vaak verschillend. Systemen zoals *Dodgeball.com* hebben als doel om kennissen samen te brengen en nieuwe mensen te leren kennen. Er zijn ook systemen die zich op de zakenwereld richten, meestal op een congres waar verschillende zakenmensen elkaar ontmoeten. Voorbeelden van systemen die zich op de zakenwereld richten zijn *nTag* en *IntelliBadge*. Verder zijn er ook systemen die enkel voor onderzoeksredenen zijn ontwikkeld, voorbeeld hiervan is het *Ubicomp Experience Project*. In volgende opsomming bespreken we kort de reeds bestaande systemen.

- ***Dodgeball.com*** : Gebruikers kunnen met een sms aangeven waar ze zich bevinden. Vrienden krijgen vervolgens een bericht op de gsm indien je dicht bij je vriend bent. Het is tevens mogelijk om friends-of-a-friend op de hoogte te stellen van je aanwezigheid.
- ***Hummingbird*** : Mobiel systeem werkende met radiofrequentie om mensen in dezelfde omgeving te laten samenwerken.
- ***nTag*** : Een pionier in de markt van elektronische badges. Profielen van gebruikers worden via de pc met infrarood naar de badge gestuurd. Als twee badges bij elkaar in de buurt komen, worden de gemeenschappelijke profieleigenschappen weergegeven. De nTag is gebaseerd op de doctoraatsthesis van Rick Borovoy [5].
- ***IntelliBadge*** : IntelliBadge gebruikt RFID om de locatie van de deelnemers te bepalen. Omdat de apparaten geen visuele output geven,

wordt een publiek scherm gebruikt waarop de informatie gepresenteerd wordt [8].

- ***SpotMe*** : SpotMe is een op Linux gebaseerd toestel dat met radiofrequentie communiceert met gelijkaardige toestellen om informatie zoals introducties aan personen, informatie over andere deelnemers en zoeken naar specifieke personen te verschaffen.
- ***Ubicomp Experience Project*** : Dit systeem maakt gebruik van RFID in combinatie met traditionele conferentiebadges. Het project maakt gebruik van gebruikersprofielen die gebruikt worden om gemeenschappelijke eigenschappen op een scherm te projecteren. [33]

2.3 Sociaal Perspectief

Uit vorige secties blijkt dat er veel mogelijkheden bestaan voor Social Proximity Sensing. Als we de interacties van alle gebruikers opslaan in een overkoepelend systeem kunnen we sociale situaties bestuderen en conclusies trekken. In het werk van o.a. Nicolai [50] staat beschreven hoe we conclusies over de connectie tussen gebruikers kunnen trekken uit sociale situaties. Nicolai focust zich op sociale gelegenheden waarbij er geen kennis is over de individuele personen. In zijn aanpak tracht hij kennis te verwerven over de navigatie van de gebruiker, bv “Is een persoon samen met anderen of passeert deze persoon toevallig de anderen?” en “Ontmoeten deze personen elkaar geregeld?”

De sociale situaties kunnen worden gecategoriseerd op basis van het soort connectie tussen beide personen. De socioloog Milgram [40] raadt aan om de categorieën als volgt in te delen :

- **Familiare personen** : Personen die men reeds kent zoals kennissen, vrienden, familie,
- **Niet-familiare personen** : Personen die men niet kent.

- Familiare vreemdeling : Een familiere vreemdeling of familiar stranger is een persoon die men geregeld tegenkomt, maar waar men nog nooit mee heeft gesproken. Een voorbeeld van een familiere vreemdeling is een persoon die men kent van op de bus of een student aan dezelfde instelling. Paulos en Goodman [36] hebben een concept ontwikkeld om familiere vreemdelingen aan de hand van hun toestel te herkennen. Ander onderzoek is geleverd door Eagle en Pentland [20], die stellen dat het mogelijk is om vrienden te herkennen op basis van Bluetooth afstanddata.

2.4 Gebruikersprofiel

Het werk van Nicolai geeft aan dat het mogelijk is om conclusies te trekken uit sociale situaties zonder hierbij een gebruikersprofiel te gebruiken. Als we echter wél gebruik maken van een gebruikersprofiel kunnen we wellicht een betere service bieden aan de gebruiker. Een gebruikersprofiel kan zeer simpel zijn, zoals bij de Lovegety [24] waar enkel het geslacht en de stemming van een persoon worden opgeslagen.

Hoe uitgebreider het profiel is, des te meer gedetailleerde en gebruikersgerichte informatie we kunnen aanbieden.

2.4.1 Richtlijnen

Aan de Universiteit van Paderborn¹ [42] heeft men een aantal vereisten opgesteld waaraan een profiel moet voldoen.

- Het is beter om gebruik te maken van verschillende profielen dan wel één groot, complex profiel;
- Eenvoudige syntax;
- Volgens standaarden;
- Mogelijkheid om afhankelijkheden tussen profielen te modeleren;

¹Paderborn University

- Mogelijkheid om makkelijk en efficiënt opties aan te passen;
- Beschrijvend en algemeen genoeg om alles te beschrijven;
- Mogelijkheid om dynamische wijzigingen in de profielen toe te staan;
- Kleine bestandsgrootte voor apparaten en netwerken met beperkte middelen.

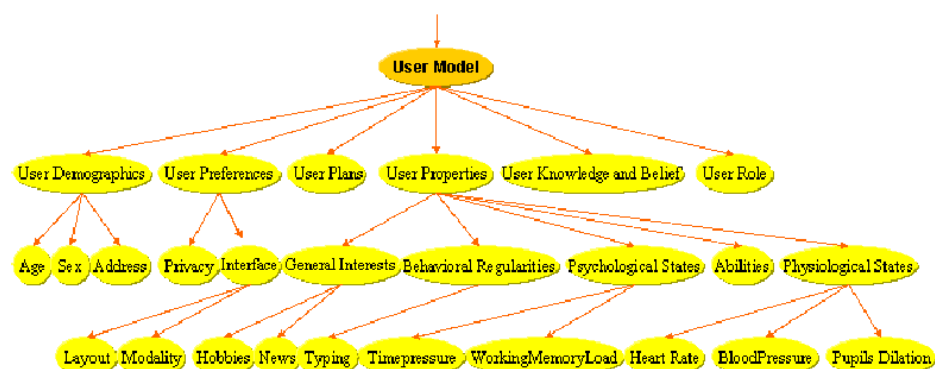
Verder geven ze aan dat er een keuze gemaakt moet worden tussen eenvoudige profielen met een intelligente evaluatie in de applicatie of complexere, zelf-lerende profielen.

2.4.2 UserML

Standaardisatie van gebruikersprofielen is van groot belang als we willen dat gebruikersprofielen hun opmars kunnen maken in de échte wereld. UserML [23] is een poging om de standaard van de gebruikersprofielen te worden. Jon Orwant formuleert in zijn Doppelgängerproject de nood aan zo'n standaard gebruikersprofiel als volgt : “*We need a protocol for encoding information about users, that any given user modeling system should be able to benefit from others and that user models should follow you around.*” [35]. UserML splitst de mogelijke informatie omtrent een gebruiker op in verschillende categorieën. Dit profiel wordt in een XML-bestand opgeslagen omdat XML makkelijk in een internetomgeving gebruikt kan worden. De verschillende categorieën worden geïllustreerd in figuur 2.3.

2.4.3 IHHI

In de context van Implicit Human-Human Interaction (IHHI) [49] is een gebruikersprofiel omschreven dat verder bouwt op de UserML en beter aansluit bij hetgeen we willen bereiken. Het gebruikersprofiel dat in IHHI gebruikt wordt, is ook in XML beschreven en kan dus via een XSLT transformatie naar een UserML notatie worden omgezet om zo eventueel aan andere



Figuur 2.3: Een taxonomie van een gebruikersmodel om mee te starten.

systemen gekoppeld te worden.

Het gebruikersprofiel van IHHI wordt opgedeeld in vier categorieën :

- Identiteit : Bevat gebruikersgegevens zoals de gebruikersnaam, geboortedatum, adres, ...
- Werk : Informatie over de tewerkstelling van de gebruiker.
- Liefde : Bijvoorbeeld de burgerlijke staat van de gebruiker.
- Vrije tijd : De wijze waarop een gebruiker zijn vrije tijd besteedt.

Voor ieder profielement wordt zowel informatie over het profielement op zich als informatie over de waarde van het profielement meegegeven. De informatie over het profielement op zich wordt onderverdeeld in :

- Auxiliary : De relatie tussen het predikaat en het onderwerp.
- Predicate : Het bestandsdeel van het statement.
- Range : Het bereik van de mogelijke waarden voor het profielement.

De informatie over de waarde van het profielement wordt opgedeeld in

:

- De eigenlijke waarde.
- De starttijd van de huidige waarde.
- De duratie van de geldigheid van de waarde.
- Een waarde die het vertrouwenslevel aangeeft.

Een voorbeeld van een gebruikersprofiel dat in IHHI gebruikt wordt is te vinden in figuur 2.4.

```
<?xml version="1.0" encoding="utf-16" ?>
- <userprofile>
- <identity>
- <situationalstatement>
  <auxiliary>hasProperty.600100</auxiliary>
  <predicate>givenname.800410</predicate>
  <range>Text.640002</range>
  <value>jan</value>
  <start>4/05/2006 19:03:49</start>
  <duration>Inherent.520002</duration>
  <detail>medium</detail>
</situationalstatement>
</identity>
- <work>
- <situationalstatement>
  <auxiliary>hasKnowledge.600120</auxiliary>
  <predicate>UbiquitousComputing.120586</predicate>
  <range>TerribleExcellent.640230</range>
  <value>Excellent</value>
  <start>4/05/2006 19:04:29</start>
  <duration>Year.520020</duration>
  <detail>low</detail>
</situationalstatement>
</work>
<love />
<freetime />
</userprofile>
```

Figuur 2.4: Voorbeeld van een gebruiksprofiel dat door IHHI gebruikt wordt.

2.4.4 FOAF-profiel

Friend of a friend (FOAF) is een term die gebruikt wordt om naar iemand te verwijzen die men niet goed kent - letterlijk, een vriend van een vriend.

Foaf is een project om sociale netwerken te modelleren zodat deze informatie door computers gelezen kan worden. Het FOAF-project is gestart door Libby Miller en Dan Brickley. Een FOAF model's hart is een schema om relaties tussen personen te beschrijven. Verder kunnen ook andere attributen zoals naam, geslacht en interesses in een FOAF-model worden beschreven. Om linken tussen de profielen toe te staan moet ieder record een unieke identifier voor iedere vriend bevatten. Voorbeelden van zo'n unieke identifier zijn SHA1 checksum van hun e-mailadres, een Jabber ID of een url van hun website of weblog ([52], [13] en [12]).

FOAF is gebaseerd op RDF en staat toe om informatie uit te wisselen tussen verschillende computer systemen. Een voorbeeld van een FAOF model staat hieronder in figuur 2.5 uitgewerkt.

2.4.5 Andere mogelijkheden

Gebruikersprofielen kunnen ook op andere manieren worden samengesteld. Steffen [45] gebruikt in het MobiSoft project bijvoorbeeld RDF voor de beschrijving van een gebruikersprofiel, Cavalluzzi [2] maakt gebruik van een Mobile User Profile (MUP), maar ook andere gebruikersprofielen zijn mogelijk.

2.5 Intelligente Profielen

Gebruikersprofielen zijn zeer handig en worden vaak voor zowel commerciële toepassingen als voor educatieve doeleinden gebruikt. Er wordt echter weinig aandacht geschonken aan het feit dat een gebruiker vaak veel informatie moet invullen en dat het systeem voldoende gebruiksvriendelijk moet zijn. Om de gebruiksvriendelijkheid te verhogen kan er gebruik gemaakt worden van een *agent*. Er zijn twee soorten agents, de *Interface Agents* en de *Autonomous Agents*. De Interface Agents [44] helpen de gebruiker bij het nemen

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <foaf:Person>
    <foaf:name>Yannick Van de Weerd</foaf:name>
    <foaf:mbox rdf:resource=
      "mailto:yannick.vandeweerd@student.uhasselt.be" />
    <foaf:homepage rdf:resource="http://yannick.kokima.be" />
    <foaf:nick>yannickv</foaf:nick>
    <foaf:depiction rdf:resource=
      "http://yannick.kokima.be/aus_img_small.jpg" />
    <foaf:interest>
      <rdf:Description rdf:about=
        "http://www.wikimedia.org" rdfs:label="Wikipedia" />
    </foaf:interest>
    <foaf:knows>
      <foaf:Person>
        <foaf:name>Kristof Ghys</foaf:name>
      </foaf:Person>
    </foaf:knows>
  </foaf:Person>
</rdf:RDF>

```

Figuur 2.5: Voorbeeld RDF bestand.

van beslissingen terwijl de Autonomous Agents volledig autonoom werken en beslissingen kunnen nemen in plaats van de gebruiker. Het belang van agenten die zowel Autonomous Agent als Interface Agent zijn is aangetoond door Lieberman [46].

De Apt Decision Agent van Lieberman [46] is een Autonomous Agent die vastgoedsuggesties aan gebruikers doet aan de hand van het gebruikersprofiel. Er was vraag naar zo'n techniek omdat gebruikers het vaak moeilijk vinden om precies te beschrijven wat ze willen hebben. De mening die mensen over een bepaald iets hebben verandert vaak gedurende de exploratie van de mogelijkheden en hetgeen beschikbaar is. De agent past het gebruikersprofiel aan door de interacties van de gebruiker in rekening te brengen. Lieberman gaat ervan uit dat iedere interactie een bedoeling heeft. In het

kort uitgelegd werkt de agent als volgt: eerst wordt er een basis profiel opgesteld door de gebruiker, vervolgens worden er aan de hand van dit profiel een aantal suggesties door het systeem gedaan. Aan de hand van de keuzes die de gebruiker maakt, wordt het profiel geüpdatet. Waarom kiest de gebruiker niet de eerste mogelijkheid maar de tweede? Wat is het verschil tussen de eerste en de tweede keuze?

Het gebruik van profielen die zelflerend zijn, kan ook gebruikt worden om de gebruiker te helpen bij het zoeken naar informatie op het internet. Billsus, Hilbert en Maynes-Aminzade [14] hebben bijvoorbeeld zo'n suggestiesysteem ontwikkeld.

2.6 Privacy and Trust

In een systeem waarin gebruikersgegevens worden uitgewisseld, moeten we letten op de privacy van de gebruiker. Om geslaagde matchings uit te voeren moet er voldoende informatie omtrent de gebruikers beschikbaar zijn. De vraag is echter in welke mate we deze informatie vrijgeven aan andere gebruikers van het systeem.

In het algemeen kunnen we stellen dat gebruikers er weinig problemen mee hebben om persoonlijke informatie mee te delen aan het systeem. Een studie van Ackerman [32] heeft uitgewezen dat slechts 17 % geen persoonlijke informatie meedeelt aan een systeem, 56 % heeft problemen met het geven van specifieke informatie en 27 % heeft (bijna) nergens problemen mee. Een studie van Spiekermann [41] is zelfs nog optimistischer.

We moeten daarom extra opletten met de informatie die we publiekelijk zichtbaar maken daar de gebruiker er zich niet steeds van bewust is dat sommige informatie de gebruiker in gevaar kan brengen. Een inbreker die van een gebruiker weet waar hij woont en wanneer hij op vakantie is, is een voorbeeld van gevaarlijke informatie. Een ander voorbeeld is iemand met slechte intenties die weet waar een jong meisje naar school gaat en welk traject ze moet volgen om naar huis te gaan. Het is meteen duidelijk dat er veel mogelijkheden zijn om informatie te misbruiken.

De privacy is sterk gelinkt aan het vertrouwen dat een gebruiker in een andere gebruiker stelt. In deze paper gebruiken we de term 'Trust' zoals deze gedefinieerd is door Deutsch [31]. Deutsch definieert Trust als :

Definitie 1. *Trust is a bet about the future contingent actions of the others.*

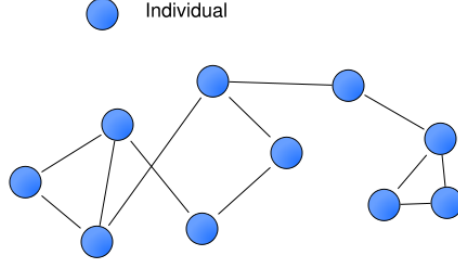
Het vertrouwen dat een gebruiker in een andere gebruiker stelt, is helaas geen simpele formule die een score teruggeeft. Aan het werk van Marsh [39] wordt vaak gerefereerd in het onderzoeksdomein over trust. Marsh zijn model is echter zeer theoretisch en complex en wordt dus bijna nooit geïmplementeerd.

Een andere wijze om een score te berekenen is om gebruik te maken van de gelijkenissen tussen twee gebruikers. De link tussen vertrouwen en gelijkenissen van gebruikers is uitgebreid onderzocht en er blijkt een grote correlatie te bestaan tussen beide factoren. Burgess & Wallin (1943) [18], Newcomb (1961) [47] en Byrne (1961 en 1971) [9] [10] tonen aan dat er *een* positieve relatie tussen vertrouwen en gelijkenissen van het profiel is. Ziegler & Golbeck [34] gaan een stap verder en tonen aan dat er een *sterke* relatie is tussen beide factoren.

Bij Implicit Human-Human Interaction (IHHI) wordt het privacyprofiel opgesplitst in verschillende modellen die overeenkomen met het gebruikersprofiel. De unie van al deze modellen vormt het privacyprofiel van een gebruiker. Om de privacy van IHHI nader te bestuderen bespreken we kort hoe de profielen zijn opgebouwd.

$$\mathcal{P} = \{P_{identity}, P_{work}, P_{love}, P_{freetime}\}$$

De contacten of vrienden (h) van de gebruiker krijgen een trust level (t) toegekend ($t \in \{Never, None, Low, Medium, High\}$). Om te voorkomen dat de gebruiker bij ieder contact afzonderlijk deze waarde moet instellen, is de mogelijkheid voorzien om gebruikers in groepen (g) op te slaan en een trust level (t) aan een volledige groep toe te kennen. Omdat een trust



Figuur 2.6: Voorbeeld van een sociaalnetwerkstructuur.

level voor een bepaald contact of voor een groep van contacten kan variëren afhankelijk van de context is er ook een context afhankelijk statement (c). Een set van groepen wordt aangeduid met het symbool G en een set met de context afhankelijke statements wordt aangeduid met C . Als we dit in formulevorm omzetten krijgen we :

$$\forall P \in \mathcal{P} : P = \{G, C\} = \{\bigcup_i \{(g_i, t_{gi})\}, \bigcup_j \{(c_j, t_{cj})\}\} : g_i = \{h_1, \dots, h_n\}$$

Een voorbeeld kan meer duidelijkheid scheppen in de formule. Als voorbeeld nemen we een gebruiker die op zijn werk alle werkgerelateerde informatie deelt met zijn collega's en minder informatie deelt met zijn vrienden. Op vakantie wil hij echter met niemand informatie delen. Dit omschrijven we als volgt :

$$P_{work} = \{G_{work}, C_{work}\}, G_{work} \supseteq \{(colleagues, High), (friends, Low)\}, - \\ C_{work} \supseteq \{(onHoliday, Never)\}$$

2.7 Social Network Sites

De term, Social Network, werd als eerste in gebruik genomen door J. A. Barnes in 1954 [1] [54]. Een sociaal netwerk kan best omschreven worden als een graaf waarbij de knopen gebruikers voorstellen en de bogen aangeven dat de gebruikers elkaar kennen. Op figuur 2.6 is dit grafisch weergegeven.

Als inleiding geven we enkele interessante stellingen die helpen bij het vormen van een idee over social networks.

Het netwerk van één gebruiker is beperkt tot 150 contacten volgens Dunbar's Number [51] [17]. Een ander interessant gegeven is de theorie van de "Six Degrees of Separation". Deze theorie stelt dat een ieder op onze planeet, via een netwerk met zes niveaus, met elkaar verbonden is. Met andere woorden wil dit zeggen dat iedereen op onze planeet met elkaar verbonden is via een netwerk, met maximaal vijf tussenschakels. Deze hypothese is voor het eerst geopperd door de Hongaarse schrijver Karinthy Frigyes. Het concept is gebaseerd op het idee dat het aantal kenniskringen exponentieel toeneemt met het aantal schakels in de netwerkketen. Op deze stelling komen we later in sectie 4.4 op terug.

Een sociaal netwerk kan verschillende vormen aannemen. Aan één zijde hebben we de kleine, sterk gekoppelde netwerken en aan de andere zijde hebben we de grote, zwak gekoppelde netwerken. Een "open" netwerk (groot netwerk met zwakke koppelingen) biedt de opportuniteit om nieuwe gebruikers en ideeën te ontdekken.

Een sociaal netwerk heeft tal van toepassingen, in deze paper richten we ons op het Internet Social Network. Deze toepassing is voor het eerst commercieel in gebruik genomen in 1995 onder de naam Classmates.com. Hierna volgden al snel andere gelijkaardige projecten, voorbeelden zijn SixDegrees in 1997, Friendster in 2002, Orkut in 2004 en MySpace in 2005. Een volledig overzicht van de belangrijkste social network sites kan men raadplegen op Wikipedia ².

Een Social Network site kan zich op verschillende doelpublieken richten. De meeste netwerken zijn gericht op het samenbrengen van personen voor gezamenlijke, persoonlijke activiteiten zoals vriendschap of een relatie. Voorbeelden zijn lnm.eu, match.com, rendez-vous.be, redbox.be, asl.be, Deze sites laten de gebruiker een profiel aanmaken om vervolgens deze

²http://en.wikipedia.org/wiki/List_of_social_networking_websites

informatie op een persoonlijke pagina te presenteren aan de andere netwerkgebruikers.

In de zakenwereld wordt echter stilaan ook meer en meer gebruik gemaakt van social network sites [26]. Een voorbeeld van een site die zich volledig op de bedrijfsmensen richt is join2grow ³. Andere sites passen hun structuur aan om uitkomst te kunnen bieden aan de groeiende vraag voor zakenprofielen, voorbeeld hiervan is Orkut ⁴ die een apart gebruikersprofiel heeft dat zich richt op de zakenwereld.

2.8 Conclusie

In dit hoofdstuk hebben we de verschillende mogelijkheden om een Profile Matching systeem te ontwikkelen onderzocht. Voor de Social Proximity Sensing maken we in onze implementatie gebruik van een Bluetooth apparaat voor de detectie van apparaten die zich in de buurt bevinden. Eens de detectie gedaan is, worden de gebruikersprofielen, gedefinieerd door IHHI en gebaseerd op UserML, uitgewisseld over het Wifi van de apparaten. De Profile Matching gebeurt op een eigen geïmplementeerd systeem waar we niet nader op ingaan daar dit een verdere studie vergt. Om het systeem te testen maken we gebruik van een Social Network Site die alle gebruikers over het internet bundelt.

³www.join2grow.biz

⁴www.orkut.com

Hoofdstuk 3

User Profile

Inhoudsopgave

3.1	Situering	24
3.2	Profile Designer	26
3.3	Soorten variabelen	26
3.4	User Profile Richtlijnen	27
3.4.1	Gebruiksgemak	28
3.4.2	Dubbelzinnige informatie	30
3.4.3	Standaardwaarden	30
3.4.4	Verplichte velden	31
3.4.5	Wijzigingen toelaten	31
3.4.6	Exacte data en aanliggende profielitems	32
3.4.7	Ordening	33
3.4.8	Belangrijke elementen	34
3.5	Context Awareness	35
3.6	Conclusie	37

In dit hoofdstuk trachten we een tot een aantal richtlijnen te komen die helpen bij het vormen van een gebruikersprofiel, User Profile genoemd.

Hiervoor kijken we naar reeds bestaande implementaties uit andere deeldomeinen zodat we ons gebruikersprofiel beter kunnen definiëren.

Vervolgens maken we een onderscheid tussen de *soorten variabelen* om deze te gebruiken bij het vormen van richtlijnen voor een User Profile. Tevens trachten we *Context Awareness* toe te voegen aan het profiel.

Uiteindelijk komen we tot een User Profile dat de gebruiker toestaat om zichzelf tot in detail te beschrijven. Daarenboven is het profiel van de gebruiker verwerkbaar door een computersysteem zodat deze profielen digitaal beheerbaar en verwerkbaar zijn.

3.1 Situering

Profile Matching is reeds in vele toepassingen aanwezig. Vaak merkt de gebruiker dit niet, toch treedt hij hiermee in contact. Steeds meer toepassingen gaan een Profile Matching toepassen om een persoonlijk voorstel te maken. Uit deze brede waaier van mogelijkheden zullen we ons moeten toespitsen op één specifiek deeldomein, namelijk het matchen van personen, User Profile Matching genoemd.

Het is zeker nuttig om de andere perspectieven van Profile Matching te bekijken, om zo onze algemene kennis over Profile Matching te verbreden. Een verduidelijking over 'Profile Matching' en 'User Profile Matching' geven we in volgende definities :

Definitie 2. *Een profiel is een beschrijving van een object. Hoe uitgebreider een profiel is, hoe meer informatie beschreven kan worden.*

Definitie 3. *Profile Matching definiëren we hier als zijnde een techniek om profielen met elkaar te vergelijken waarbij rekening gehouden wordt met het zoekprofiel van de gebruiker. Op dit zoekprofiel komen we later terug.*

Definitie 4. *User Profile Matching beperkt zich tot het vergelijken van profielen die een persoon beschrijven. User Profile Matching is hetzelfde als Profile Matching maar richt zich enkel op het vergelijken van gebruikersprofielen en niet op algemeen beschrijvende profielen.*

In onderstaande paragrafen geven we voorbeelden van waar profile matching toegepast wordt. Deze voorbeelden zijn vaak niet meteen zichtbaar voor de gebruiker terwijl hij er toch gebruik van maakt.

Muziekaanbevelingsdiensten vinden hun weg naar het grote publiek. Meestal verschijnen deze diensten in de vorm van extensies voor de bekende mediaspelers. Een voorbeeld hiervan is [28].

Tegenwoordig is het normaal dat een e-commerce site aanbevelingen kan doen over andere interessante producten voor de gebruiker. Deze aanbevelingen zijn persoonlijk en vaak gebaseerd op het koopverleden van deze persoon en hetgeen andere gebruikers gekocht hebben.

Er zijn tal van applicaties die toelaten om op een meer sociale manier het web te verkennen. Deze applicaties verschijnen meestal in de vorm van plugins of extensies voor bekende webbrowsers. Voorbeelden zijn Dai.sy [11] en Ratepoint [38]. Bij Dai.sy kunnen gebruikers die zich op éénzelfde webpagina bevinden discussiëren over een gekozen onderwerp. Ratepoint gaat hier nog een stap verder in. Gebruikers die éénzelfde pagina bezoeken kunnen discussiëren over een gekozen onderwerp, maar kunnen ook een score aan een site geven. Indien een gebruiker deze pagina bezoekt zal hij een score over deze pagina te zien krijgen, die gebaseerd is op de score die andere gebruikers reeds gegeven hebben. Daarenboven zal de score van een website, die een gebruiker te zien krijgt, afhangen van zijn persoonlijke voorkeuren.

“RatePoint changes the way you look at ratings by personalizing them based on your likes and dislikes, connecting you with like-minded people who

share the same interests.”

Dai.sy bevat dus een User Profile Matching. Deze is echter beperkt tot het gemeenschappelijk bekijken van éézelfde webpagina. De User Profile Matching van Ratepoint houdt ook rekening met persoonlijke voorkeuren.

3.2 Profile Designer

Een User Profile moet door vrijwel iedere gebruiker gebruikt kunnen worden. Er zijn echter aanpassingen in een User Profile die meer inzicht en kennis vragen dan we van een standaard eindgebruiker kunnen verwachten. Deze complexere aanpassingen worden door een Profile Designer gedaan.

Definitie 5. *Een Profile Designer is een persoon waarvan we meer kennis verwachten dan een standaard eindgebruiker. Deze designer is verantwoordelijk voor het aanpassen en beheren van de profielstructuur.*

3.3 Soorten variabelen

Vooraleer we starten met de bespreking over user profile richtlijnen definiëren we nominale en ordinale variabelen. De definities van deze soorten variabelen zijn geformuleerd door Dirk Speelman [29].

Definitie 6. *een nominale variabele is een variabele waarvan de mogelijke waarden voor geen enkele rekenkundige operatie of vergelijking in aanmerking komen.*

Een voorbeeld van een nominale variabele is “kleur van de ogen” met mogelijke waarden “blauw”, “groen”, “bruin”, Deze eigenschap kunnen we niet optellen (“blauw” + “groen”) of vermenigvuldigen (“blauw” $\times$ “groen”). Het is tevens onbelangrijk dat “groen” in de opsomming na “blauw” staat.

Definitie 7. *een ordinale variabele is een variabele waarvan de waarden een inherente orderrelatie hebben, maar zonder de garantie dat rekenkundige operaties zoals tellen of vermenigvuldigen zinvol zijn.*

Voorbeeld van een ordinale variabele is een variabele die aangeeft hoe goed iemand in iets is. Zo kan “tenniskwaliteit” een variabele zijn met “slecht”, “goed” en “zeer goed” als waarden.

Bij deze eigenschap heeft de optelling of vermenigvuldiging van waarden ook geen nut. De volgorde is hier echter in tegenstelling tot het vorige voorbeeld wel van belang; het is belangrijk dat “slecht” voor “goed” in de volgorde staat.

3.4 User Profile Richtlijnen

Een basisvereiste om een degelijke profile matching uit te voeren, is dat de profielen van de gebruikers voldoende uitgebreid en correct zijn. In de volgende paragrafen bespreken we waaraan een goede profielstructuur moet voldoen.

In een profiel moeten we een onderscheid maken tussen de verschillende profielelementen en hun bijbehorende keuzes. In figuur 3.1 is “Java kennis” een profielelement terwijl “Goed”, “Zeer goed” en “Uitstekend” de bijbehorende keuzes voor “Java” zijn.

Definitie 8. *Een profielelement is een eigenschap die in het profiel opgenomen wordt. Een profielelement bevat vaak een aantal keuzeopties waarbij de gebruiker aangeeft hoe deze eigenschap naar zichzelf gelinkt moet worden.*

Definitie 9. *Een profielitem zijn de keuzes behorende tot een profielelement. Uit deze profielitems moet de gebruiker een keuze maken zodat een profielelement het best de gebruiker beschrijft.*

Voor het opstellen van een User Profile hebben we een aantal richtlijnen opgesteld. Deze richtlijnen zijn eigen bevindingen, tenzij anders aangegeven



Figuur 3.1: Voorbeeld van een profielement.

bij de richtlijn. Indien een User Profile zich aan deze richtlijnen houdt, kunnen we een Profile Matching systeem voorstellen dat met het User Profile kan werken. Op deze matching komen we in hoofdstuk 4 terug.

Onderstaande richtlijnen zullen we in volgende secties in detail bespreken.

- Gebruiksgemak;
- Opletten met dubbelzinnige informatie;
- Standaardwaarden mijden;
- Verplichte velden mijden;
- Verwelkom (late) wijzigingen;
- Maak een verschil tussen exacte en niet-exacte data;
- Orden de profielementen;
- Aangeven van belangrijkheid van profielementen.

3.4.1 Gebruiksgemak

De profielementen moeten makkelijk en intuïtief aanpasbaar zijn. Bij de afwezigheid van een groot gebruiksgemak zal de gebruiker ervan afzien om veel informatie aan te passen. Dit is tevens een richtlijn die Robbie Schaefer geeft in “Fuzzy Evaluation of User Profiles” [42]. Andere richtlijnen van

The image shows a dark-themed user profile form. It is organized into four distinct sections, each with a yellow title: 'uiterlijk' (appearance), 'programming', 'Going out', and 'More profile items'. The 'uiterlijk' section contains two items: 'kleur ogen :' with a dropdown menu showing 'zwart' and 'lengte :' with a dropdown menu showing '1 m60'. The 'programming' section contains one item: 'Java :' with a dropdown menu showing 'Goed'. The 'Going out' section contains one item: 'cafe :' with a dropdown menu showing 'ilem'. Each of these four items has a red 'X' mark to its right, indicating that the data is missing or incomplete. At the bottom of the form, there is a green plus icon and the text 'More profile items'.

Figuur 3.2: Voorbeeld van een profiel waarbij profielelementen gegroepeerd kunnen worden en waar niet-ingevulde profielelementen niet worden getoond.

Robbie Schaefer werden reeds besproken in hoofdstuk 2.

Een voorbeeld om het gebruiksgemak te verhogen is het groeperen van profielelementen. Dit omdat bij een groot aantal profielelementen het overzicht snel verloren zal gaan. Door elementen te groeperen, verkrijgen we terug een beter overzicht op de data. Een andere mogelijkheid die eventueel gecombineerd kan gaan met het groeperen van profielelementen is om de niet-ingevulde profielelementen niet te laten zien aan de gebruiker. De gebruiker moet dan wel in staat zijn om gewenste profielelementen toe te voegen. Het grote voordeel van deze tweede oplossing is dat de gebruiker niet overladen wordt met overbodige informatie.

Een combinatie van beide oplossingen is geïllustreerd in figuur 3.2. De profielelementen 'kleur ogen' en 'lengte' zijn bevat in de groep 'uiterlijk' terwijl 'Java' en 'café' ieder in een aparte groep zitten. In dit profiel zijn enkel de ingevulde profielelementen weergegeven. Andere profielelementen kunnen worden toegevoegd door op 'More profile items' te klikken.

3.4.2 Dubbelzinnige informatie

Dubbelzinnige profielementen kunnen de betrouwbaarheid van profile matching sterk ondermijnen. Als verschillende gebruikers een element anders interpreteren kunnen fouten optreden bij de matching. Indien een gebruiker A een element als zijnde x interpreteert en gebruiker B interpreteert hetzelfde element als y , dan zal gebruiker A ontevreden zijn over de matching resultaten indien gebruiker B in de lijst staat. Een voorbeeld kan een “computerkennis” element in een profiel zijn. Mogelijk vindt een secretaresse haar computerkennis zeer goed daar ze vlot kan werken met diverse bureau-applicaties (tekstverwerker, databasebeheerprogramma, ...). Een professor in de computerwetenschappen zal in zijn zoektocht naar mensen met een zeer goede kennis van de computer ontevreden zijn over de matching indien de secretaresse op de eerste plaats staat. Dit omdat de professor een ander idee heeft over het “computerkennis”-element.

Dit probleem is te voorkomen door een profielement in een juiste groep te plaatsen. Indien de groepsnaam een indicatie geeft van hoe het profielement geïnterpreteerd moet worden, kan dit al meer duidelijkheid scheppen voor de gebruiker.

In het voorbeeld van “computerkennis” zou dit beter opgesplitst kunnen worden in verschillende profielementen zoals “Programmeerervaring”, “Bureau-applicatiekennis”, “Software-Engineeringkennis”, Algemeen kunnen we stellen dat een grotere granulariteit van de verstrekte informatie minder kans op dubbelzinnigheden heeft.

3.4.3 Standaardwaarden

Door een standaardwaarde toe te kennen aan een profiel, kan het zijn dat iemand die niets persoonlijk heeft ingegeven, toch gemiddeld scoort bij een profile matching. Dit is een ongewenst neveneffect en moet dus vermeden worden. Een voorbeeld is een profielement dat de kennis over een specifiek deeldomein beschrijft (zoals “kennis over software engineering”). Het is onwaarschijnlijk dat iemand zijn kennis hierover “goed” is, terwijl “goed”

vaak een standaardwaarde van dit profielement zal zijn. De kennis van een leek zal eerder “slecht” of “zeer slecht” zijn. Dit is echter ook in twijfel te trekken daar de persoon in kwestie misschien een natuurlijke aanleg heeft voor “software engineering” maar deze kennis nooit getest heeft. In dit geval weet de gebruiker dus gewoonweg niet wat zijn kennis hieromtrent is. Dit is de reden dat een standaardwaarde in veel gevallen ongewenst is.

De oplossing die we hiervoor verkiezen, is dezelfde als besproken in sectie 3.4.1. Door profielementen die niet ingevuld zijn weg te laten en niet op te nemen in verdere berekeningen hebben we geen behoefte aan een standaardwaarde.

3.4.4 Verplichte velden

Door de gebruiker te verplichten om een lange lijst met profielitems in te vullen, zal deze geënerveerd raken en velden foutief invullen om tijd te besparen. Dit euvel kan voorkomen worden door de gebruiker te laten kiezen welke items hij/zij wenst in te vullen. Let wel dat we niet vervallen in voorgaande fout door standaardwaarden toe te kennen aan de niet-ingevulde items.

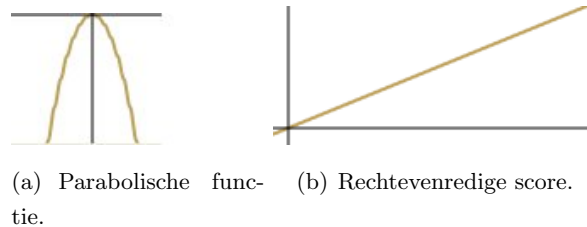
Door in de berekeningen rekening te houden met de ingevulde en de niet-ingevulde profielementen hebben we geen nood aan deze standaardwaarden.

3.4.5 Wijzigingen toelaten

Het moet mogelijk zijn om profiel informatie aan te passen indien nodig. De kans is bestaande dat een gebruiker na verloop van tijd andere waarden wenst in te vullen voor de profielitems. Als deze mogelijkheid niet bestaat, of als de grens om aan te passen te groot is¹, zal de gebruiker ervan afzien om het profiel te wijzigen en zal het profiel bijgevolg foute data bevatten.

¹Dit kan doordat de gebruiker moet betalen, een ingewikkelde paginastructuur, bevestigingsmails, ...

3.4.6 Exacte data en aanliggende profielitems



Figuur 3.3: Grafieken die aangeven hoezeer aanliggende profielitems meetellen in de berekening.

Bij het vergelijken van profielementen moet er een onderscheid worden gemaakt tussen exacte elementen of elementen die met de vergelijkingsoperatoren $=$, $<$, $>$, $\leq$ en $\geq$ met elkaar kunnen worden vergeleken. We illustreren dit a.d.h.v. een voorbeeld. De grootte van een persoon is een exact element, indien een gebruiker op zoek is naar iemand met een bepaalde grootte is een groter iemand niet noodzakelijk beter. Terwijl een score op een vaardigheid wel de eigenschap bevat dat een hogere score vaak beter is. Dit soort variabelen waarbij een exacte matching vereist is, worden in o.a. de statistiek beschreven als *nominale variabelen* (zie definitie 6).

De andere profielitems behorende tot een profielement zijn *ordinale variabelen* (zie definitie 7). Deze profielitems staan dus wel toe om items met elkaar te vergelijken.

Om aan te geven hoezeer de aanliggende profielitems, die sterk gelijken op de 'perfecte' match, meegerekend moeten worden, maken we gebruik van een door de gebruiker gecreëerde grafiek.

Dit illustreren we best met een voorbeeld waarbij we opnieuw verwijzen naar figuur 3.1. In het voorbeeld kan er voor het "Java" element gekozen worden voor "Goed", "Zeer goed" of "Uitstekend". Wat gebeurt er met de matchingscore indien de gewenste keuze "Zeer goed" is, maar de gebruiker

waarmee gematched wordt, heeft “Goed” aangeklikt?

In onze implementatie maken we gebruik van een grafiek die de verhouding tussen de score en de omliggende profielitems aangeeft. In deze grafiek kan worden aangegeven dat een aangrenzende waarde (“Goed” en “Uitstekend in het voorbeeld) voor 50 % meetelt, of dat deze waarden voor 0 % meetellen. In de grafiek kan ook worden aangegeven dat ”Uitstekend“ beter is dan ”Zeer goed“ zodat indien er gezocht wordt naar iemand met een ”Zeer goede“ kennis, iemand met een ”Uitstekende“ kennis ook wordt opgenomen in de resultatenlijst.

In figuur 3.3 worden twee mogelijke functies weergegeven. Grafiek 3.3(a) geeft een parabolische functie weer. Indien we deze parabolische functie vertalen naar een score, zullen profielitems die verder verwijderd zijn van de ’ideale’ keuze een lagere score krijgen dan de items die dicht bij de ’ideale’ keuze liggen.

Grafiek 3.3(b) geeft daarentegen een rechtevenredig verband aan tussen de afstand en de score. Dit houdt in dat profielitems die ver verwijderd zijn van de ideale keuze een hogere score zullen ontvangen dan de items die dicht bij de ’ideale’ score liggen. Merk op dat in dit geval de ’ideale’ keuze de minst goede score ontvangt! Deze grafiek zal dus gebruikt worden om negatieve eigenschappen aan te geven.

Van een eindgebruiker mogen we niet verwachten dat deze probleemloos een grafiek kan interpreteren en opstellen. Dit is best een taak die we toekennen aan de Profile Designer.

3.4.7 Ordening

De data moeten geordend kunnen worden. Een alfabetische volgorde of een volgorde op datum van aanmaak is vaak niet de beste oplossing. De belangrijkste profiel groepen moeten bovenaan gezet worden terwijl de minder belangrijke data onderaan de weergave gezet moeten worden. Een voorbeeld

The screenshot shows a dark-themed user profile editor. At the top, the group 'uiterlijk' is highlighted in yellow. Below it, there are two items: 'kleurogen' and 'lengte'. Each item has a dropdown menu set to 'uiterlijk', a 'Combobox' dropdown, and two checkboxes: 'exact' (checked) and 'more is better' (unchecked). To the right of each item is a text input field containing '4*x' for 'kleurogen' and 'x*x' for 'lengte', followed by a 'getGraph' button. Below each item is a 'show values' button. A 'new item' button is also present. The 'programming' group is shown below, with the item 'Java' having similar settings and a 'x*x' value. Blue arrows above the group names indicate they can be moved. Red 'X' marks are visible next to the 'getGraph' buttons and the 'programming' group name.

Figuur 3.4: Ordening van de profielgroepen.

staat in figuur 3.4 waar "uiterlijk" door middel van de blauwe pijltjes onder de groep "programming" gezet kan worden. Bij het verplaatsen van een groep worden uiteraard alle bijbehorende elementen meegenomen. Doordat we een volgorde bepalen op de groepen, zijn de groepen tevens *ordinale variabelen*.

Ook dit is een taak die we best aan de profile designer toekennen.

3.4.8 Belangrijke elementen

Bepaalde eigenschappen wegen zwaarder door dan andere eigenschappen. De gebruiker moet dit kunnen aangeven zodat een profile matching dichter bij zijn persoonlijke voorkeuren aanleunt. Door ordinale waarden die een belangrijkheid aanduiden te koppelen aan de profielelementen, kan een gebruiker aangeven hoe belangrijk hij/zij dit element vindt.

Een voorbeeld hiervan staat in figuur 3.5. In het voorbeeld kan de gebruiker aangeven hoe belangrijk hij een bepaald profielelement vindt. Één of

The image shows a dark-themed user interface for profile evaluation. It contains two sections. The first section is titled 'uiterlijk' in yellow. Below it, there is a label 'lengte :', a dropdown menu showing '1 m30', and a row of five stars with the first star highlighted in yellow. The second section is titled 'programming' in yellow. Below it, there is a label 'Java :', a dropdown menu showing 'Goed', and a row of five stars with the first four stars highlighted in yellow.

Figuur 3.5: Het geven van een score aan een profielement door een star-rating mechanisme.

twee ster(ren) koppelt een negatieve score aan dit element, drie sterren zijn neutraal en meer sterren (vier of vijf) geven een positief resultaat. In dit geval is een lengte van 1m30 een zwaar minpunt bij de matching, terwijl een goede kennis van Java een pluspunt is.

3.5 Context Awareness

Een profiel beschrijving van een persoon is vaak slechts geldig in bepaalde omstandigheden. Zo kan het voorkomen dat een gebruiker in een bepaalde context liever minder informatie vrijgeeft dan in andere contexten.

Definitie 10. *In de computerwetenschappen wordt de term "Context Aware" gebruikt om te refereren naar computers die een omgeving kunnen interpreteren en hierop vervolgens kunnen reageren. Toestellen kunnen informatie bevatten waarin beschreven staat hoe ze moeten reageren in bepaalde omstandigheden. De term "Context-awareness" wordt geïntroduceerd door Shilit [3].*

Definitie 11. *Context aware toestellen trachten veronderstellingen te maken over de huidige situatie van de gebruiker. Dey [16] definieert context als "om het even welke informatie die gebruikt kan worden om de situatie te karakteriseren".*

Om dit concept beter te begrijpen illustreren we het concept "Context Awareness" met behulp van een voorbeeld. Stel dat een werknemer van een

bepaalde firma op vakantie is. Het is dan aannemelijk dat deze persoon niet wenst gestoord te worden door andere personen met werkgerelateerde vragen. Het moet dus mogelijk zijn om tijdelijk de informatie over het werk niet te verspreiden. In ons voorbeeld van de werknemer die op vakantie is kan het zijn dat deze persoon zijn hobbies en algemene interesses wel wenst te verspreiden terwijl hij hiervan het nut niet inziet als hij aan het werk is.

Uit het voorbeeld blijkt dat er naargelang de context andere profiel informatie vrijgegeven moet worden. Het profiel moet aan deze vereiste kunnen voldoen.

Om aan de voorgestelde eis te voldoen beschikt een gebruikersprofiel over vier categorieën die naargelang de context gekozen worden. De categorieën hebben we overgenomen uit het werk over Digital Aura door Carl Bruninx [7]. De categorieën sommen we even op :

- Identity
- Work
- Love
- Freetime

Door gebruik te maken van een toestel dat de context van de gebruiker kan bepalen kan een categorie uit het profiel van gebruiker gekozen worden zodat de vrijgegeven informatie beter aansluit met de wens van de gebruiker. Carl Bruninx maakt in zijn werk gebruik van een PDA, maar een gsm is ook reeds succesvol gebruikt in andere onderzoeken zoals dat van Nathan Norfleet Eagle [19]. Het is van belang dat we ervanuit kunnen gaan dat de aanwezigheid van het toestel impliceert dat de eigenaar van het toestel zich in de buurt van het toestel bevindt.

3.6 Conclusie

In vorige sectie werden een aantal richtlijnen gepresenteerd die helpen bij het vormen van een goede User Profile structuur. Er wordt aangeraden om al deze richtlijnen te gebruiken bij het opstellen van een profiel. Verder wordt het belang van Context Awareness aangetoond waarbij een profiel zich aanpast naargelang de context waarin de gebruiker zich bevindt.

Hoofdstuk 4

User Profile Matching

Inhoudsopgave

4.1	User Profile Matching Structuur	39
4.2	Analyse en opsplitsing van het probleem	39
4.3	Basic Matching	42
4.4	Foaf Matching	45
4.5	Trusted Matching	48
4.6	Conclusie	51

In het vorige hoofdstuk hebben we besproken waaraan een goede profielbeschrijving moet voldoen. In dit hoofdstuk gebruiken we een User Profile dat voldoet aan alle richtlijnen die besproken werden.

User Profile Matching is een systeem waarbij gebruikers profielen (User Profiles) met elkaar worden vergeleken en dat vervolgens een score teruggeeft die de mate van overeenkomst tussen beide profielen aangeeft. Deze score zullen we de matching score noemen.

Om zulk systeem te maken, bespreken we eerst kort de algemene structuur om vervolgens verder op de verschillende componenten van het Profile

Matching systeem in te gaan. De verschillende componenten zijn Basic Matching, Foaf Matching en Trusted Matching.

4.1 User Profile Matching Structuur

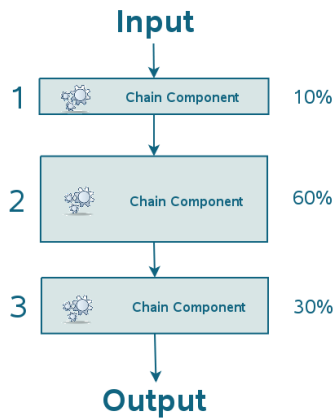
Profielmatching systemen bestaan in alle mogelijke smaken, gaande van zeer simpele systemen tot gesofistikeerde systemen die complexe algoritmen gebruiken om een goed resultaat te bekomen. Voorbeeld van een simpele profile matching is de Lovegety [24], ontstaan in Japan. De Lovegety heeft een succesvolle profile match indien beide personen van een verschillend geslacht zijn en beide personen in dezelfde stemming (vrolijk, verdrietig, ...) zijn. Een voorbeeld van een ingewikkelde en zeer theoretische profile matching is de profile matching van Marsh [39] die vanwege zijn zeer theoretisch model zelden in praktijk wordt gebracht.

De profile matching die we hier bespreken, tracht een goede, persoonlijke uitkomst te leveren die voldoende flexibiliteit biedt en toepasbaar is in praktijkgerichte toepassingen. Om deze profile matching te testen maken we gebruik van een webapplicatie (zie sectie 5.1) waar we de theorie meteen in praktijk brengen. Zo zijn we zeker dat het algoritme dat we aanbrengen niet zuiver theoretisch is, zoals Marsh zijn werk, maar vlot zijn weg naar de praktijk kan vinden.

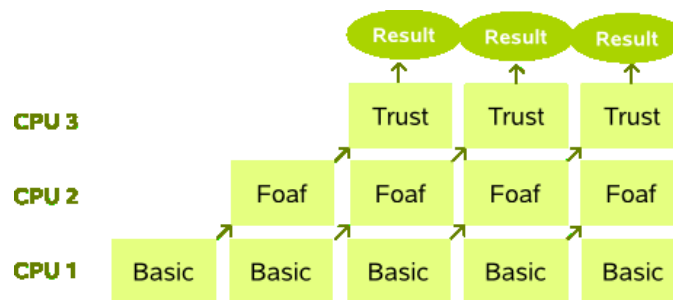
4.2 Analyse en opsplitsing van het probleem

Een uitgebreide profile matching is moeilijk in één geheel te analyseren. Daarom splitsen we de profile matching op in verschillende componenten die we vervolgens in detail bestuderen.

Om de verschillende componenten tot één geheel te verweven, maken we gebruik van een kettingstructuur waarin de componenten de schakels van de ketting voorstellen. Indien we een input aan het begin van de ketting



Figuur 4.1: Eenvoudige voorstelling van een chain waarin verschillende componenten doorlopen worden om een uitkomst te bekomen.



Figuur 4.2: Opsplitsing van berekening chain over verschillende CPU's.

invoeren, zal iedere component apart een score berekenen. Om een totaal-score te berekenen vermenigvuldigen we ieder resultaat van de componenten met zijn percentage. Dit is geïllustreerd in figuur 4.1 waarin de percentages respectievelijk 10, 60 en 30 % zijn.

Een kettingstructuur wordt o.a. in de computergraphics [15] gebruikt om van wereldcoördinaten naar schermcoördinaten over te gaan. Deze omzetting gebeurt in verschillende stappen die ieder apart berekend kunnen worden. In de computergraphics wordt deze eigenschap vaak benut om de

berekeningen van de verschillende stukken op te splitsen naar verschillende computers en zo de uitvoertijd aanzienlijk te verminderen.

Het opsplitsen en verdelen van de rekenkracht die vereist is om een matching score te berekenen is een eigenschap die nuttig kan zijn indien het gebruikers-netwerk uitbreidt en er veel matching scores op korte tijd berekend moeten kunnen worden. In figuur 4.2 wordt geïllustreerd hoe deze opsplitsing in een snellere werking kan resulteren.

Zoals aangegeven wensen we een flexibele structuur die makkelijk uitbreidbaar en aanpasbaar is. Door te werken met een kettingstructuur kunnen we deze vereisten verwezenlijken. Aangezien iedere component (schakel in de ketting) afzonderlijk werkt, kunnen we eenvoudig componenten toevoegen en/of verwijderen uit de ketting.

Een bijkomend voordeel is dat we aan iedere component een waarde kunnen koppelen die aangeeft hoe zwaar deze component doorweegt in de totale score zoals aangegeven in de vorige paragraaf en figuur 4.1.

In onze implementatie hebben we een basic matching component, een trusted matching component en een foaf component. Deze componenten bespreken we in de komende paragrafen in detail. Deze componenten hebben we tijdens de literatuurstudie opgesteld. Steeds als we een nood zagen aan een nieuwe matching mogelijkheid, hebben we getracht deze toe te voegen aan een bestaande component en, indien nodig, een nieuwe component aan te maken. Moesten er toch nog nieuwe profiel matching vereisten opduiken, dan kunnen we deze eenvoudig aan de huidige structuur toevoegen.

Om een score uit de ketting van componenten te bekomen, maken we gebruik van onderstaande formule waarbij w_i het gewicht van component i voorstelt, m_i de matchscore van component i voorstelt en n het aantal componenten. De formule is als volgt :

$$output = \sum_{i=1}^n w_i m_i$$

4.3 Basic Matching

Zoals aangegeven bestaat onze User Profile Matching uit die componenten. De Basic Matching Component zullen we hier als eerste bespreken. Deze Basic Matching component vergelijkt profielementen op basis van de gekozen profielitems behorende tot profielementen. Deze component houdt geen rekening met andere factoren zoals betrouwbaarheid of foaf-afstand. Betrouwbaarheid en foaf-afstand zullen in rep sectie 4.5 en sectie 4.4. Vooraleer we deze component in detail toelichten, definiëren we eerst enkele termen.

Definitie 12. *De waarderingsscore is een score die een gebruiker aan een profielement geeft om aan te geven hoe belangrijk hij/zij deze eigenschap vindt.*

Een voorbeeld is te vinden in figuur 4.4 waar de gebruiker zijn waardering kan uiten door een aantal sterren te selecteren.

Een tweede term die we graag even nader toelichten is de user created graph.

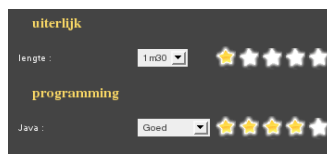
Definitie 13. *Een User Created Graph is een grafiek die aangeeft hoezeer aanliggende profielitems in een profielement worden meegerekend.*

De voordelen van deze grafiek zijn reeds besproken in 3.4.6.

Deze component bepaalt een score door een profielement te vergelijken met een overeenkomstig profielement. De Basic Matching Component kunnen we omzetten naar een wiskundig model. Hiervoor maken we gebruik van diverse subformules om het overzicht te bewaren. Deze formules bespreken we in onderstaande paragrafen ieder apart.



Figuur 4.3: Een profielement met bijbehorende profielitems waaruit gekozen kan worden.



Figuur 4.4: Een profielement met bijbehorende waarderingscore.

In de implementatie kan een gebruiker een score aan een profielement koppelen. Deze score geeft aan hoe belangrijk ze een bepaalde eigenschap achten. Om deze score (1, 2, 3, 4 of 5) om te zetten naar een bruikbare waarde gebruiken we onderstaande formule :

$$score_{rating} = \begin{cases} -3 & \text{als } x = 1 \\ -1 & \text{als } x = 2 \\ 0 & \text{als } x = 3 \\ 1 & \text{als } x = 4 \\ 3 & \text{als } x = 5 \end{cases}$$

Hieruit volgt dat indien een gebruiker slechts één ster aanvinkt, dit profielement met een waarde van -3 meetelt in de berekening.

De waarde behorende tot een score hebben we zelf bepaald. Mogelijk geven andere waarden betere resultaten in bepaalde omstandigheden. Het is echter wel belangrijk dat de waarde behorende tot een score steeds hoger

is ten opzichte van de waarde behorende tot een lagere score. Verder moet er aan gedacht worden om rekening te houden met deze verandering in de verdere berekeningen. Zo is de maximale score, indien we gebruik maken van de voorgestelde waarden, 3. Indien deze waarden veranderd zijn, veranderd ook de maximale score. Op deze maximale score komen we later in deze sectie terug.

De gebruiker heeft de mogelijkheid om een functie (*userCreatedGraph*) aan een profielement te koppelen. Deze functie geeft aan hoezeer de waarde van dat profielement daalt indien er een andere optie uit de mogelijkheden is gekozen.

In figuur 4.3 is de afstand tussen "blauw" en "bruin" bijvoorbeeld één terwijl de afstand tussen "blauw" en "groen-blauwig" drie is. Deze waarde wordt ingevuld in de relevantiegrafiek en deze uitkomst zal vervolgens met de waarderingsscore vermenigvuldigd worden om zo een score voor dit profielement te bepalen. De formule hiervoor is :

$$score_{graph} = \frac{\maxProfileElementVal}{userCreatedGraph(x)}$$

waarbij *maxProfileElementVal* een maximaal mogelijke waarde voor deze profielitems is.

Merk op dat de *score_{graph}* genormaliseerd is tussen [0,1]. Deze eigenschap zullen we later nodig hebben om een maximale waarde voor een profiel te kunnen bepalen.

Om uit de twee voorgaande formules een formule voor een profielement te bekomen, volstaat het om beide formules met elkaar te vermenigvuldigen.

$$score_{profileElement} = score_{rating} \times score_{graph}$$

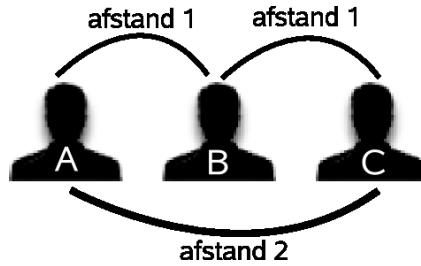
De score voor het gehele profiel wordt bekomen door de maximale score, *maxScore*, te delen door de sommatie van alle scores van de profielementen.

menten n . De maximale score $maxScore$ is het aantal profielementen, n , vermenigvuldigd met drie, omdat drie de maximale waarde is die een profielement kan krijgen. Dit maakt dat drie de maximale waarde is die een profielement kan krijgen omdat de waarderingsscore maximaal drie is en de $score_{graph}$ maximaal één is. Daar deze twee scores met elkaar vermenigvuldigd moeten worden, kunnen we maximaal een score van drie bekomen voor een profielement.

$$score = \frac{maxScore}{\sum_{i=0}^n score_{profileElement}(i)} \quad (4.1)$$

$$= \frac{n \times 3}{\sum_{i=0}^n \{score_{rating} \times (\frac{maxProfileElementVal}{userCreatedGraph(x)})\}} \quad (4.2)$$

4.4 Foaf Matching



Figuur 4.5: Grafische representatie van de foaf-afstand tussen personen.

Onder foaf verstaan we friend-of-a-friend. Deze term werd volledig besproken in het hoofdstuk 2 in sectie 2.4.4. De foaf-afstand is het aantal tussenpersonen dat nodig is om persoon A in contact te brengen met persoon B. Indien persoon A en B elkaar kennen is de foaf-afstand één, indien persoon A een persoon B kent en persoon B kent op zijn beurt persoon C, dan is de foaf-afstand tussen persoon A en persoon C twee. Indien persoon A, persoon C rechtstreeks kent is de foaf-afstand slechts één. We veronderstellen dat indien de foaf-afstand vermeld wordt, dat dit de kortste foaf-afstand

tussen twee personen betreft. Dit wordt geïllustreerd in figuur 4.5.

Karinthy Frigyes stelt met haar "Six Degrees of Separation" dat iedereen op onze planeet met elkaar verbonden via een netwerk, met maximaal vijf tussenschakels. Milgram komt tot eenzelfde conclusie in zijn "*small world experiment*". Het experiment was als volgt geconstrueerd [53] :

1. Hoewel het experiment uit verschillende variaties bestaat, koos Milgram meestal individuen uit Omaha, Wichita en Boston om als start- of eindpunt van de ketting te dienen. Deze steden zijn gekozen omwille van hun grote afstand in de VS. Deze afstand is zowel sociaal als geografisch groot.
2. Informatie pakketten werden willekeurig verzonden naar individuen uit Omaha en Wichita. Deze pakketten bevatten brieven waarin gedetailleerd het doel van de studie wordt uitgelegd. Verder bevat het pakket basis informatie over de doelpersoon in Boston. Daarenboven bevat het pakket een schema waarop de deelnemer zijn gegevens achterlaat en deze naar de onderzoekers kan versturen.
3. Bij de ontvangst van de uitnodiging om deel te nemen, werd de ontvanger gevraagd of hij/zij persoonlijk de contactpersoon kende die in de brief wordt beschreven. Indien dit het geval was, moest de persoon het pakket aan die persoon rechtstreeks versturen.
4. Het is waarschijnlijk dat de persoon niet rechtstreeks de doelpersoon kent. In dit geval moesten de deelnemers het pakket doorgeven aan een kennis waarvan ze verwachtten dat deze kennis de doelpersoon wel kent. Vervolgens werd gevraagd om hun naam aan het schema toe te voegen en hun gegevens naar de onderzoekers te sturen.
5. Indien het pakket zijn eindbestemming in Boston bereikt had, konden de onderzoekers bepalen hoeveel tussenpersonen er nodig waren om van start- tot doel-persoon te geraken. Voor pakketten die hun eindbestemming niet bereikte, kon het breekpunt in de ketting worden

bepaald doordat iedere deelnemer zijn gegevens naar de onderzoekers doorstuurde.

Uit deze test bleek dat de gemiddelde afstand tussen start- en doelpersoon 5,5 tot 6 was. Dit cijfer komt overeen met hetgeen Karinthy Frigyes reeds vermeldde in haar "Six Degrees of Separation".

Psychologen hebben gedurende tientallen jaren bestudeerd welke eigenschappen personen opmerken en zoeken in mogelijke partners. Brehm et al. [6] en Fisher [22] hebben een overzicht gepresenteerd met de diverse factoren die bepalen wat iemand attractief vindt. Hieruit bleek dat foaf-afstand en familiariteit de belangrijkste factoren bij een dating matching zijn [48].

Deze component geeft een score aan de hand van de foaf-afstand tussen de twee gematchte gebruikers en bevat dus geen onverwachte complicaties. Een foaf-afstand van maximaal twee is goed. Een grotere afstand geeft problemen omdat er dan niemand is die beide partijen aan elkaar kan voorstellen. Daarom wordt er nooit verder dan twee diep in de graaf gekeken [30].

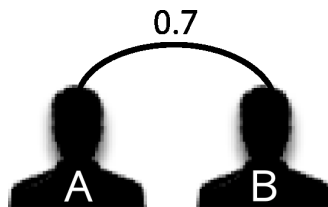
Afhankelijk van de foaf-afstand zal deze component een hoge of lage waarde als resultaat bekomen. Als we deze matchingcomponent in formulevorm zetten verkrijgen we volgende formule :

$$score = \begin{cases} 0 & \text{als } x \geq 3 \\ 0,3 & \text{als } x = 2 \\ 1 & \text{als } x = 1 \end{cases}$$

De arbitraire waarden 0, 0.3 en 1 zijn door ons gekozen daar deze goede resultaten gaven. Deze waarden zijn echter vrij te kiezen. Het is wel belangrijk dat de waarden omgekeerd evenredig zijn met de foaf-afstand zodat er bij een grotere foaf-afstand, een kleinere score gegeven wordt. Wij hebben ervoor gekozen om de waarden niet recht-evenredig te verdelen. Dit omdat

indien er via een kennis moet worden gewerkt om een persoon te introduceren, dit beduidend minder gewenst is. Daarom is de waarde behorende tot een foaf-afstand van 2 slechts een waarde van 0.3. Het is tevens mogelijk om een waarde te koppelen aan foaf-afstanden van meer dan 3, dit is echter niet aan te raden omdat er dan niemand is die beide partijen aan elkaar kan voorstellen.

4.5 Trusted Matching



Figuur 4.6: Illustratie van de trust berekening tussen twee personen die elkaar rechtstreeks kennen.

Betrouwbaarheid is een belangrijk goed in de maatschappij. Een systeem dat de mogelijkheid biedt om rekening te houden met iemand zijn betrouwbaarheid heeft dus steeds een belangrijke stap voor op een systeem dat deze functie niet ondersteunt.

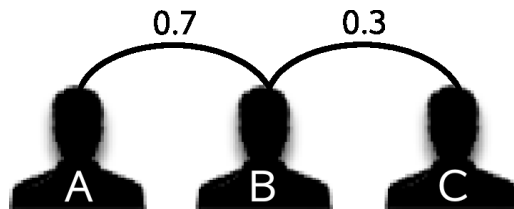
De betrouwbaarheid van een contactpersoon wordt door de gebruiker zelf bepaald. De gebruiker heeft de mogelijkheid om een betrouwbaarheidsscore te koppelen aan een andere gebruiker. Indien alle gebruikers een score geven aan al zijn of haar contactpersonen kunnen we een betrouwbaarheidsscore berekenen voor iemand in het netwerk die de gebruiker niet rechtstreeks kent. De betrouwbaarheidsscore werkt in één richting. Persoon A kan persoon B zeer goed vertrouwen terwijl persoon B persoon A niet vertrouwt.

Deze component zoekt naar de maximale trustwaarde tussen twee personen. We hebben ervoor geopteerd om de maximale waarde te nemen daar

een gebruiker niet snel geneigd is om een onbekend persoon sterk te vertrouwen. We kunnen dus veronderstellen dat indien een gebruiker een sterk vertrouwen in iemand heeft, dat dit zijn reden heeft. Let wel: dit is een vermoeden; we kunnen onmogelijk foute betrouwbaarheidwaarden uitsluiten.

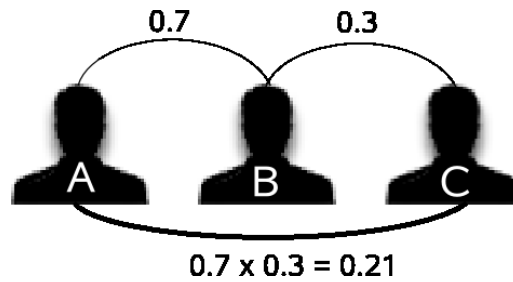
De berekeningen die deze component uitvoert kunnen best uitgelegd worden aan de hand van verschillende voorbeelden. Om de trustscore tussen twee personen te berekenen die elkaar kennen, wordt simpelweg de score genomen die de gebruiker aan deze contactpersoon gekoppeld heeft. Dit wordt geïllustreerd in figuur 4.6.

Indien de gebruikers elkaar niet rechtstreeks kennen, wordt de trustscore berekent door de diverse trustscores met elkaar te vermenigvuldigen. Dit leggen we ook best uit met een voorbeeld. Stel dat persoon A persoon B kent en persoon B kent op zijn beurt persoon C. De trustscore wordt berekent door de trustscore die persoon A aan persoon B geeft te vermenigvuldigen met de trustscore die persoon B aan persoon C heeft gegeven. In formulevorm dus $trustscore1 \times trustscore2$. In het voorbeeld van figuur 4.7 moeten we dus 0.7 met 0.3 vermenigvuldigen, zoals in figuur 4.8 geïllustreerd wordt.



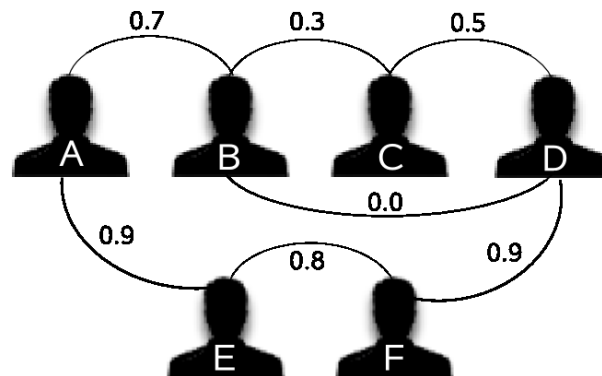
Figuur 4.7: Illustratie van de trust berekening waarbij gebruik gemaakt word van een tussenpersoon.

Indien persoon A persoon C wél rechtstreeks kent, wordt de hoogste trustscore genomen. Als we er echter voor geopteerd hadden om de laagste trustscore te nemen zou dit nadelige gevolgen kunnen hebben. Een gebruiker kan een andere gebruiker saboteren door een nulscore te geven in de



Figuur 4.8: Resultaat van een trust-score berekening waarbij gebruik gemaakt wordt van een tussenpersoon.

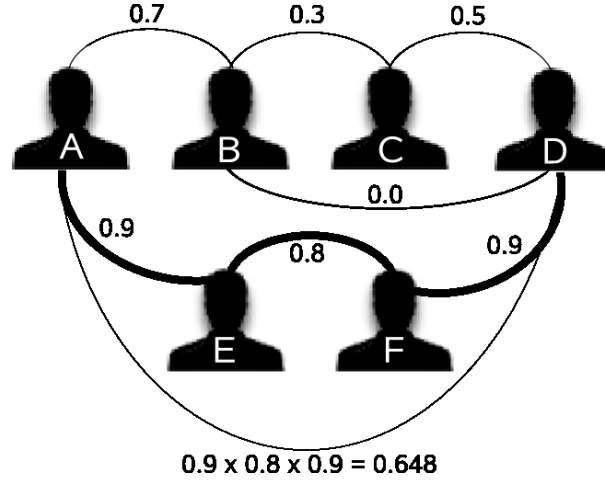
wetenschap dat deze persoon bijgevolg steeds als onbetrouwbaar beschouwd wordt in het systeem. In figuur 4.9 kan gebruiker B, gebruiker D *altijd* benadelen door een 0 score te geven. Merk op dat voor iedere gebruiker een pad naar gebruiker D bestaat waarbij gebruiker B opgenomen is, waardoor er steeds een totale score van 0 bekomen wordt. Door het pad te nemen tussen de begin- en eindgebruiker waarbij we de hoogste trust-score verkrijgen, bekomen we het in vet aangeduide pad in figuur 4.10. De totale trust-score voor dit pad is $0.8 \times 0.9 \times 0.8 = 0.648$.



Figuur 4.9: Illustratie van de trust berekening waarbij een gebruiker een andere gebruiker een nul-score geeft.

Deze component kunnen we ook in formulevorm zetten.

Om de contactrelatie te beschrijven, gebruiken we een $contact(i,j)$ predicaat.



Figuur 4.10: Illustratie van de trust berekening waarbij verschillende paden mogelijk zijn.

Dit predicaat geeft aan dat persoon i , persoon j kent. Om de betrouwbaarheid tussen persoon i en j te beschrijven gebruiken we een $trust(i, j)$ functie die aangeeft hoezeer persoon i , persoon j vertrouwt. Stel de verzameling

$$(i_1, \dots, i_n) : contact(i_j, i_{j+1})$$

dan kunnen we de betrouwbaarheidsscore tussen twee personen als volgt berekenen :

$$\prod_{j=1}^n trust(i_j, i_{j+1})$$

Zoals reeds aangegeven aan het begin van deze sectie is een ondersteuning van Trust sterk gewenst in Profile Matching. Door een eenvoudige toevoeging kunnen we de waarde van de Profile Matching aanzienlijk verhogen.

4.6 Conclusie

We herhalen kort de gemaakte conclusies van vorige sectie.

Om een matching score te bekomen doorlopen we een kettingstructuur waar-

bij iedere schakel in de ketting een component voorstelt. Om een score uit de ketting van componenten te bekomen, maken we gebruik van onderstaande formule waarbij w_i het gewicht van component i voorstelt, m_i de matchscore van component i voorstelt en n het aantal componenten. De formule is als volgt :

$$output = \sum_{i=1}^n w_i m_i$$

Als componenten hebben we voor een Basic Matching, Foaf Matching en Trust Matching gekozen. De Basic Matching component vergelijkt profielementen op basis van de gekozen profielitems behorende tot profielementen. De score voor het gehele profiel wordt bekomen door de maximale score, $maxScore$, te delen door de sommatie van alle scores van de profielementen n . De formule die we hiervoor bekomen zijn is als volgt :

$$score = \frac{maxScore}{\sum_{i=0}^n score_{profileElement}(i)} \quad (4.3)$$

$$= \frac{n \times 3}{\sum_{i=0}^n \{score_{rating} \times (\frac{maxProfileElementVal}{userCreatedGraph(x)})\}} \quad (4.4)$$

De Foaf Matching houdt rekening met hoe goed gebruikers elkaar kennen. De formule die we hiervoor bekomen hebben is als volgt :

$$score = \begin{cases} 0 & \text{als } x \geq 3 \\ 0,3 & \text{als } x = 2 \\ 1 & \text{als } x = 1 \end{cases}$$

Tot slot hebben we de Trust Matching, die het vertrouwen dat gebruikers in elkaar hebben, mee opgenomen in de berekening van een totaal score. Deze component kunnen we ook in formulevorm zetten. Om de contactrelatie te beschrijven, gebruiken we een $contact(i,j)$ predicaat. Dit predicaat geeft aan dat persoon i , persoon j kent. Om de betrouwbaarheid tussen persoon i en j te beschrijven gebruiken we een $trust(i,j)$ functie die aangeeft hoezeer persoon i , persoon j vertrouwt. Stel de verzameling

$$(i_1, \dots, i_n) : \text{contact}(i_j, i_{j+1})$$

dan kunnen we de betrouwbaarheidsscore tussen twee personen als volgt berekenen :

$$\prod_{j=1}^n \text{trust}(i_j, i_{j+1})$$

Met deze structuur en componenten behaalden we goede subjectieve resultaten. Bovendien bieden we voldoende flexibiliteit om aanpassingen of uitbreidingen te ondersteunen.

Hoofdstuk 5

Implementatie

Inhoudsopgave

5.1	Website	55
5.1.1	Inloggen	56
5.1.2	Aanpassen van profielinformatie	56
5.1.3	Aanpassen van zoekprofiel	58
5.1.4	Aanpassen van profielementen door profile editor	58
5.1.5	Zoeken naar gebruikers	61
5.1.6	Beheren van contacten	63
5.1.7	Profielen van andere gebruikers bekijken	65
5.1.8	Exporteren contacten naar PDA applicatie	65
5.1.9	Conclusie Website	65
5.2	Serverapplicatie	65
5.2.1	Network Users & Network Graph	66
5.2.2	Selected User	68
5.2.3	Component Weights	68
5.2.4	Components Info	69
5.2.5	Calculation	69
5.2.6	Synchronizatie met website	70
5.3	Synchronisatie Server	70

In de literatuur staat veel over Profile Matching beschreven. Profile Matching is echter een onderzoeksdomein dat zeer goed in de praktijk toepasbaar is. Hieruit volgt dat een grondige testing van de theorie aan de praktijk gewenst is.

Voor deze testing hebben we een testomgeving opgesteld waarin we de gemaakte veronderstellingen kunnen toetsen aan de werkelijkheid. Deze testomgeving bestaat uit twee delen. Één deel bestaat uit een website waarop gebruikers kunnen inloggen en profielinformatie vrijgeven. Het tweede deel bestaat uit een server die voldoende feedback geeft zodat er makkelijk conclusies uit de matchingmethode kunnen worden getrokken. Dit hoofdstuk zal deze twee applicaties in detail bespreken.

5.1 Website

Om profielinformatie van de gebruikers te vergaren leek een website ons de beste keuze daar veel gebruikers reeds vertrouwd zijn met allerlei social network sites zoals www.lnm.eu, www.h5.com en www.rendez-vous.be.

Op de website is het mogelijk om het gebruikersprofiel, zoals gedefinieerd in definitie 2, aan te passen. Het is tevens mogelijk om het zoekprofiel aan te passen en contacten te beheren. We definiëren een zoekprofiel als volgt :

Definitie 14. *Een zoekprofiel is een profielbeschrijving van een persoon die interessant lijkt te zijn. Indien er een goed en uitgebreid zoekprofiel wordt opgesteld, zal de gebruiker hier profijt uit halen bij een Profile Matching.*

Definitie 15. *Contacten zijn personen die de gebruiker aan zijn contactenlijst heeft toegevoegd. Soms wordt op andere sites naar contacten verwezen onder de naam “vrienden” of “kennissen”.*

De mogelijkheden van de website sommen we even kort op. Deze mogelijkheden lichten we later in detail toe.

- Inloggen van gebruikers
- Aanpassen van profielinformatie
- Aanpassen van zoekprofiel
- Aanpassen van profielelementen door profile editor
- Zoeken naar gebruikers
- Beheren van contacten
- Profielen van andere gebruikers bekijken

5.1.1 Inloggen

Op het internet is Web 2.0 aan een grote opmars bezig. User Based Content waarbij gebruikersgerichte informatie aan de gebruiker wordt gepresenteerd, staat vaak in de kijker. Op onze website laten we de gebruiker inloggen om te weten welke gebruiker aanwezig is op het systeem. Dit om persoonlijke informatie te kunnen presenteren over o.a. de andere netwerkgebruikers.

5.1.2 Aanpassen van profielinformatie

In hoofdstuk 3 werden enkele richtlijnen voor profiel aangeboden. De richtlijnen opgesteld in dat hoofdstuk zijn veelal tot stand gekomen doordat er tijdens de implementatie problemen aan het licht kwamen die via een theoretische weg minder snel ontdekt zouden worden. We herhalen deze richtlijnen kort :

- Gebruiksgemak;
- Opletten met dubbelzinnige informatie;



The screenshot shows a dark-themed profile editing interface. It is divided into three sections, each with a title in yellow: 'uiterlijk', 'programming', and 'Going out'. Under 'uiterlijk', there are two items: 'kleur ogen :' with a dropdown menu showing 'zwart' and a red 'X' icon, and 'lengte :' with a dropdown menu showing '1 m60' and a red 'X' icon. Under 'programming', there is one item: 'Java :' with a dropdown menu showing 'Goed' and a red 'X' icon. Under 'Going out', there is one item: 'cafe :' with a dropdown menu showing 'item' and a red 'X' icon. At the bottom left, there is a green plus icon and the text 'More profile items'.

Figuur 5.1: Bewerk bestaande groepen en items en voeg groepen en items toe aan het profiel.

- Standaardwaarden mijden;
- Verplichte velden mijden;
- Verwelkom (late) wijzigingen;
- Maak een verschil tussen exacte en niet-exacte data;
- Orden de profielementen;
- Aangeven belangrijkheid van profielementen.

Het aanpassen van een profiel bespreken we adhv figuur 5.1. Bij ieder profielement kan gekozen worden uit verschillende profielitems. De gebruiker maakt een keuze uit de aangeboden profielitems waardoor het profielement gekoppeld wordt aan het gekozen profielitem.

Een profielement kan tevens verwijderd worden waardoor deze eigenschap niet meer opgenomen wordt in het profiel. Indien een gebruiker op zoek gaat naar andere gebruikers met deze eigenschap zal de gebruiker die dit profielement heeft verwijderd, niet worden opgenomen in de resultaten,

tenzij er andere eigenschappen meedoen aan de matching.

In het profiel van de gebruiker worden enkel de ingevulde profielelementen getoond, dit om het overzicht te bewaren. Profielelementen kunnen makkelijk toegevoegd worden aan het profiel. Dit moet makkelijk zijn daar een gebruiker anders minder geneigd zou zijn om profielinformatie aan te leveren. Om profielelementen aan het profiel toe te voegen dient de gebruiker in figuur 5.1 te klikken op “More Profile Items” om profielelementen aan het profiel toe te voegen.

Het is belangrijk op te merken dat een profiel element steeds aan een bepaalde categorie wordt toegewezen. Deze opdeling is merkbaar in figuur 5.2. Dit is gedaan zodat een gebruiker een bepaalde categorie kan kiezen naargenag de context waarin de gebruiker zich bevindt. De categorieën zijn : Identity, Work, Love en FreeTime. Deze zijn besproken in sectie 3.5.

5.1.3 Aanpassen van zoekprofiel

Het aanpassen van een zoekprofiel is sterk gelijkend op het aanpassen van het persoonlijke profiel. Ook hier kunnen profielelementen toegevoegd, verwijderd en aangepast worden. Dit profiel wordt door het systeem gebruikt om aanbevelingen te kunnen doen aan de gebruikers. Het is dus van groot belang dat de gebruiker dit profiel zo uitgebreid en gedetailleerd mogelijk beschrijft zodat het systeem goede aanbevelingen kan doen.

Het zoekprofiel is ook opgesplitst in categorieën zodat er naar andere profielelementen gezocht wordt naargelang de context waarin de gebruiker zich bevindt.

5.1.4 Aanpassen van profielelementen door profile editor

Een Profile Designer is een gebruiker die de profielelementen moet beheren. Deze Profile Designer wordt reeds in definitie 5 gedefinieerd. Dit deel

Identity profile

bussiness
marketshare :

Going out
cafe :

More profile items

Work profile

uiterlijk
kleur ogen :

programming
Java :

More profile items

Love profile

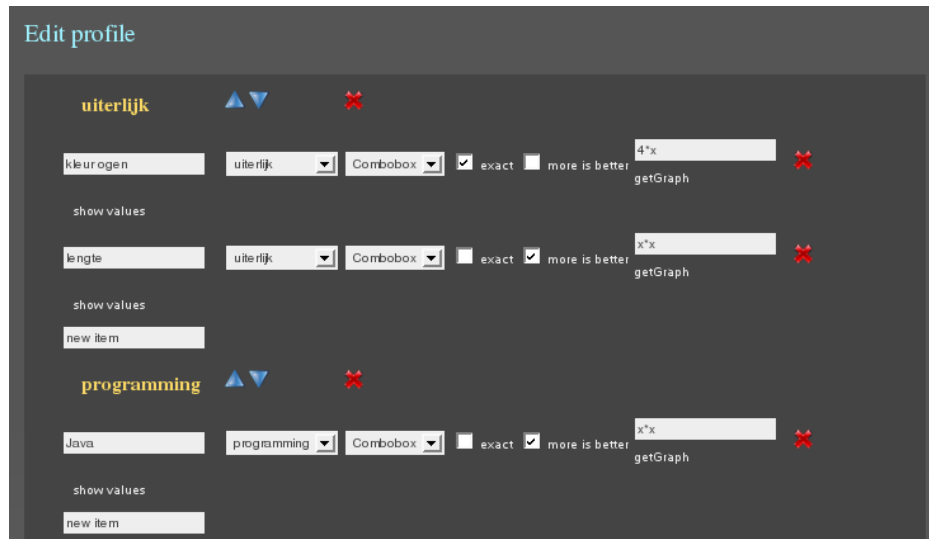
uiterlijk
lengte :

More profile items

Freetime profile

More profile items

Figuur 5.2: Profilelementen zijn opgedeeld in de categorieën Identity, Work, Love en FreeTime.



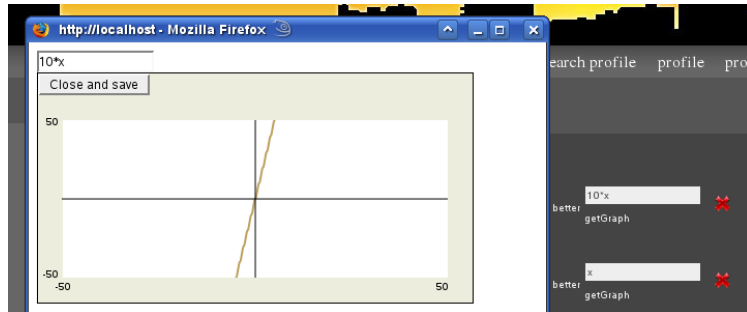
Figuur 5.3: Aanpassen van profielementen door een Profile Designer.

van de website heeft als doel om profielementen gedetailleerd te omschrijven en te beheren. Volgende mogelijkheden zijn ingewerkt in dit deel van de website :

- Naam wijzigen van profielementen
- Groeperen van profielementen
- Toevoegen, verwijderen, naam wijzigen en ordenen van groepen
- Toevoegen, verwijderen, naam wijzigen en ordenen van profielitems

Aanpassingen gemaakt door de Profile Designer zullen onmiddellijk zichtbaar zijn voor de gebruikers. De gebruikers bezitten echter niet dezelfde rechten als een Profile Designer en zijn bijgevolg beperkt in hun aanpassingsmogelijkheden.

De groepen kunnen een andere naam krijgen en kunnen verplaatst worden in de ordening. De ordening van de groepen zorgt ervoor dat de 'belangrijkste' data bovenaan de pagina staan, terwijl de minder interessante



Figuur 5.4: Grafiek editor om de relatie tussen de score en de profielitems aan te geven.

informatie onderaan staat. 'Belangrijkste' moet algemeen geïnterpreteerd worden omdat sommige groepen voor bepaalde personen belangrijker zijn dan voor andere personen. Daarom moet de profile editor rekening houden met hetgeen het meest overeenstemt met het idee dat de meeste gebruikers over de belangrijkheid van een groep hebben.

profiel-elementen bestaan uit profiel-items. De profile editor kan aan een profiel-element profiel-items toevoegen, verwijderen of van naam wijzigen. Tevens kan de volgorde van de profiel-elementen gewijzigd worden.

De profile editor kan een grafiek die de relatie tussen de score en de profiel-items aangeeft via de webapplicatie aanmaken. Het nut van zo'n grafiek werd reeds aangetoond in sectie 3.4.6. Kort komt het erop neer dat de Profile Designer door middel van een grafiek kan aangeven hoezeer aanliggende profiel-items in een profiel-element meetellen in de Profile Matching.

5.1.5 Zoeken naar gebruikers

De eindgebruiker kan via de site zoeken naar de andere gebruikers. Dit kan handig zijn indien een gebruiker snel wil weten welke gebruikers een bepaalde eigenschap bezitten.

The image shows a web application interface with a dark gray background. It contains several sections for filtering users based on different criteria. Each section has a title in a light blue font and a list of options with checkboxes. The options are preceded by a yellow arrow icon.

- Java**
 - Goed ☒
 - Zeer Goed ☒
 - Uitstekend ☐
- kleur ogen**
 - zwart ☒
 - groen ☐
 - blauw ☐
 - bruin ☐
- café**
 - item ☐
- lengte**
 - 1m60 ☐
 - 1m40 ☐
 - 1m30 ☐
 - 1m50 ☐
- marketshare**
 - big ☒
 - small ☐
 - medium ☒

At the bottom right, there is a "Search" button with a dotted border. Below it, there is a search bar containing the text "yannickv".

Figuur 5.5: Webpagina om naar andere gebruikers in het netwerk te zoeken.

Een screenshot van de zoekfunctionaliteit is te vinden in figuur 5.5. De gebruiker klikt de gewenste eigenschappen aan en klikt vervolgens op de 'Search'-knop waarna het systeem op zoek gaat naar gebruikers die aan de criteria voldoen.

Indien er meerdere profielitems van hetzelfde profielelement zijn aangevinkt, wordt naar gebruikers gezocht die aan één van de mogelijkheden voldoet. In het voorbeeld mag een gebruiker zijn 'Java'-kennis 'goed' of 'zeer goed' zijn. De gebruiker is wel verplicht zwarte ogen te hebben.

Deze zoekfunctie kunnen we abstraheren naar een logische formule. Hierbij maken we gebruik van het predicaat I_a dat *true* of *false* teruggeeft naargelang dit profielitem geselecteerd is of niet. Verder definiëren we P_i als zijnde :

$$P_i = I_1 \vee \dots \vee I_n \quad (5.1)$$

waarbij $I_1 \dots I_n$ allen profielitems van profielelement i zijn.

Het resultaat van de zoekfunctie kunnen we dan beschrijven als :

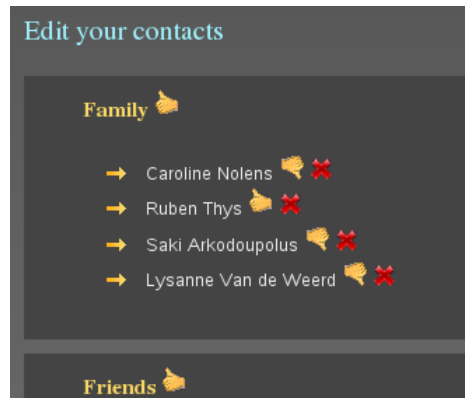
$$result = P_1 \wedge \dots \wedge P_m \quad (5.2)$$

$$= (I_{1_1} \vee \dots \vee I_{1_n}) \wedge \dots \wedge (I_{m_1} \vee \dots \vee I_{m_n}) \quad (5.3)$$

5.1.6 Beheren van contacten

Het beheren van contacten of vrienden is ook voorzien in de implementatie. Gebruikers kunnen andere netwerkgebruikers toevoegen aan hun contactenlijst en deze vervolgens groeperen in zelfgekozen groepen.

De trustwaarde die nodig is voor het berekenen van een trustwaarde in sectie 4.5 kan aan een contact gekoppeld worden door meermaals op het handje te klikken. Er zijn drie mogelijke trustwaarden (low, medium en high).



Figuur 5.6: Beheren van de contacten en toekennen van een trust score.

Een groep kan ook een default trustwaarde meekrijgen. Deze wordt op dezelfde wijze toegekend als bij een contact.

De trustwaarde is afhankelijk van de context. Er zijn vier contextmogelijkheden :

- Identity
- Work
- Freetime
- Love

Deze categorieën zijn gekozen naar analogie met de categorieën in het model beschreven in IHHI (zie sectie 2.6). Een contact of groep kan dus een andere trust waarde hebben afhankelijk van de context.

In figuur 5.6 wordt geïllustreerd hoe dit in de site wordt verwezenlijkt. Zo hebben in het voorbeeld de gebruiker rubent en de gebruiker tester een trustwaarde van 'low' terwijl de gebruiker 'tomc' een trustwaarde van 'medium' heeft.

5.1.7 Profielen van andere gebruikers bekijken

De gebruikers kunnen naar de profielen van andere gebruikers kijken. Dit kan nuttig zijn als men via de zoekfunctie interessante gebruikers ontdekt heeft en meer informatie over deze gebruiker wilt raadplegen.

5.1.8 Exporteren contacten naar PDA applicatie

Zoals reeds gesteld in de doelstelling van de thesis is het wenselijk dat er een mogelijkheid bestaat om de contacten van de website te exporteren naar een formaat dat leesbaar is door het Digital Aura framework dat opgesteld is aan het edm [49]. Op de website hebben we een knop voorzien waarop de gebruiker moet klikken, waarna de contacten door het systeem gesynchroniseerd worden. Op het synchroniseren van de contacten komen we later terug in sectie 5.3.

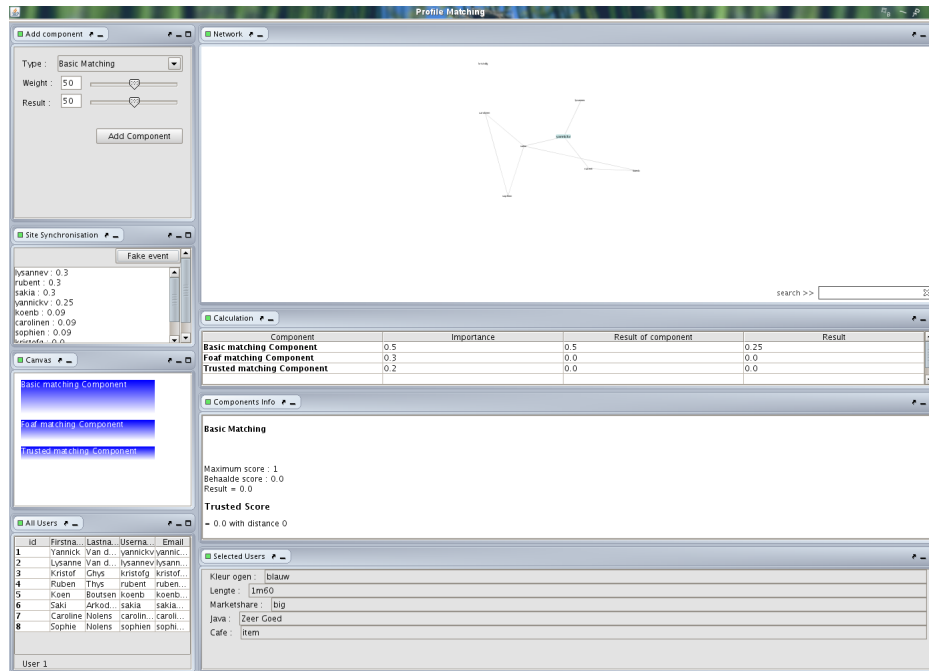
5.1.9 Conclusie Website

De implementatie van de website toont duidelijk aan dat het mogelijk is om gebruikersprofielen op te stellen die aan alle eisen voldoen zoals beschreven in het hoofdstuk 4 over User Profile Matching.

5.2 Serverapplicatie

De Profile Matching wordt op een aparte server uitgevoerd, dit om de computer van de gebruiker niet te overbelasten en om snelheidsredenen. Er is snelheidswinst daar de serverapplicatie op dezelfde server als de database draait. Hierdoor is er een kleinere overhead dan indien de serverapplicatie en de database - mogelijk ver - verspreid staan.

Een bijkomend voordeel van onze serverapplicatie is dat de applicatie een overzicht van het netwerk en de profile matching tracht te bieden. De applicatie is onderverdeeld in diverse subvensters die ieder verantwoordelijk



Figuur 5.7: Overzicht van de serverapplicatie.

zijn voor de visualisatie van een bepaald aspect. Deze vensters bespreken we ieder apart in onderstaande secties. Een overzicht van de volledige applicatie staat in figuur 5.7.

5.2.1 Network Users & Network Graph

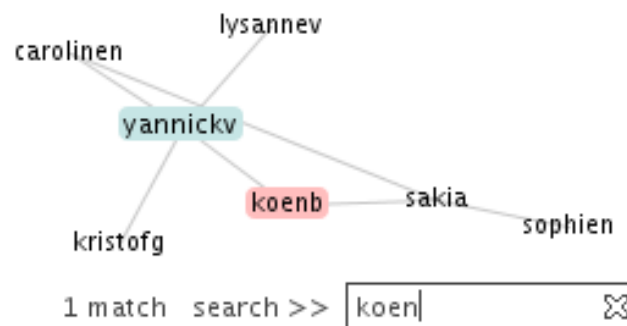
De serverapplicatie biedt een overzicht op het netwerk in zijn geheel. Een gedetailleerde weergave staat in figuur 5.8, waar iedere gebruiker van het netwerk in vermeld staat, samen met zijn naam en e-mailadres.

De relaties tussen de netwerkgebruikers wordt voorgesteld in een user graph (zie figuur 5.9). Deze graaf toont hoe gebruikers met elkaar geconnecteerd zijn. Deze informatie is bijzonder handig bij het analyseren van de foaf-component (zie sectie 4.4, wiens berekeningen berust zijn op de foaf-

All Users

id	Firstname	Lastname	Username	Email
1	Yannick	Van de We...	yannickv	yannickv@digitalaur...
2	Lysanne	Van de We...	lysannev	lysannev@digitalaur...
3	Kristof	Ghys	kristofg	kristofg@digitalaur...
4	Ruben	Thys	rubent	rubent@digitalaura...
5	Koen	Boutsen	koenb	koenb@digitalaura...
6	Saki	Arkodoupol...	sakia	sakia@digitalaura.be
7	Caroline	Nolens	carolinen	carolinen@digitalau...
8	Sophie	Nolens	sophien	sophien@digitalaur...

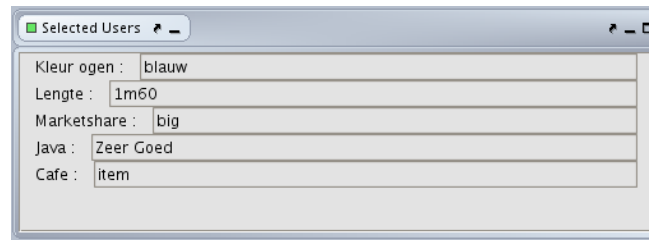
Figuur 5.8: Overzicht van alle netwerkgebruikers.



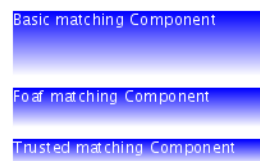
Figuur 5.9: Graaf van alle netwerkgebruikers.

afstand). De graaf is gevisualiseerd door middel van een *radial layout*. Het voordeel van deze visualisatie is dat het meteen duidelijk is hoe de gebruikers met elkaar geconnecteerd zijn, en welke gebruikers een sleutelrol spelen bij het verbinden van verschillende groepen. Bij het selecteren van een nieuwe node als root node zal de overgang geanimeerd verlopen zodat de gebruiker zo min mogelijk verward raakt. Voor een uitgebreide bespreking van deze techniek verwijzen we naar [55] en [27].

Wij hebben voor de implementatie gebruik gemaakt van de Vizster [25] toolkit. We hebben voor deze toolkit geopteerd wegens de flexibiliteit en prestatie van het pakket. Deze toolkit is tevens open-source, er zijn dus geen beperkingen op het gebruik hiervan.



Figuur 5.10: Profiel van de geselecteerde gebruiker.



Figuur 5.11: Illustratie van de gewichten die aan de componenten gekoppeld zijn.

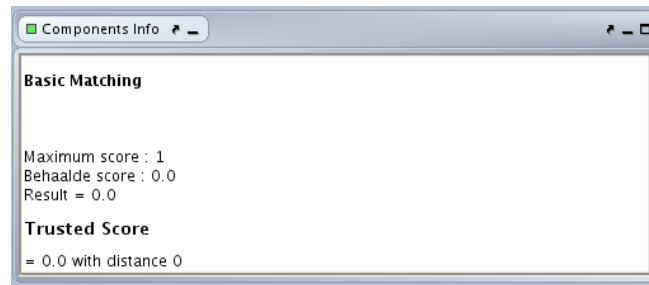
5.2.2 Selected User

Dit venster toont gedetailleerde profielinformatie over de geselecteerde gebruiker. Deze informatie is handig bij het analyseren van de basic matching component, beschreven in 4.3.

5.2.3 Component Weights

Ieder component heeft een bepaald gewicht om een component veel of weinig in de totaalberekening te laten meewegen. Deze berekening wordt reeds besproken in sectie 4.2. Deze gewichten worden visueel voorgesteld in de GUI.

Voor deze implementatie hebben we gebruik gemaakt van Piccolo [37]. Deze toolkit staat toe om snel en makkelijk figuren op een canvas te tekenen.



Figuur 5.12: Overzicht van de berekeningen benodigd voor een component.

The screenshot shows a window titled 'Calculation' containing a table with the following data:

Component	Importance	Result of component	Result
Basic matching Component	0.5	0.5	0.25
Foaf matching Component	0.3	0.0	0.0
Trusted matching Component	0.2	0.0	0.0
			0.25

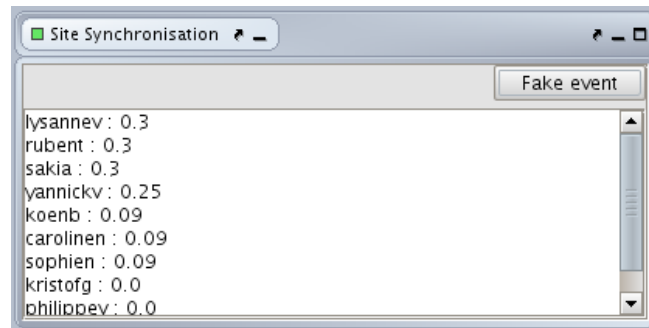
Figuur 5.13: Uitkomst van iedere component, vermenigvuldigd met een gewicht om tot een eindresultaat te komen.

5.2.4 Components Info

Het Components Info venster heeft als doel de berekeningen gemaakt door de verschillende componenten overzichtelijk te presenteren. Iedere component kan naar dit venster informatie uitschrijven. Dit venster is uitermate nuttig om componenten hun werking te testen en eventueel te debuggen.

5.2.5 Calculation

Een overzicht van de berekeningen die nodig zijn om een totaalscore te bekomen. Hier wordt aangegeven wat de inbreng van iedere component is en hoe het uiteindelijke resultaat bekomen wordt uit het samenvoegen van de diverse componentscores.



Figuur 5.14: Informatie over de aanvragen via de website.

5.2.6 Synchronizatie met website

Daar we spreken over een serverapplicatie is het duidelijk dat er clientconnecties gemaakt moeten kunnen worden. De gemaakte connecties worden in dit venster weergegeven.

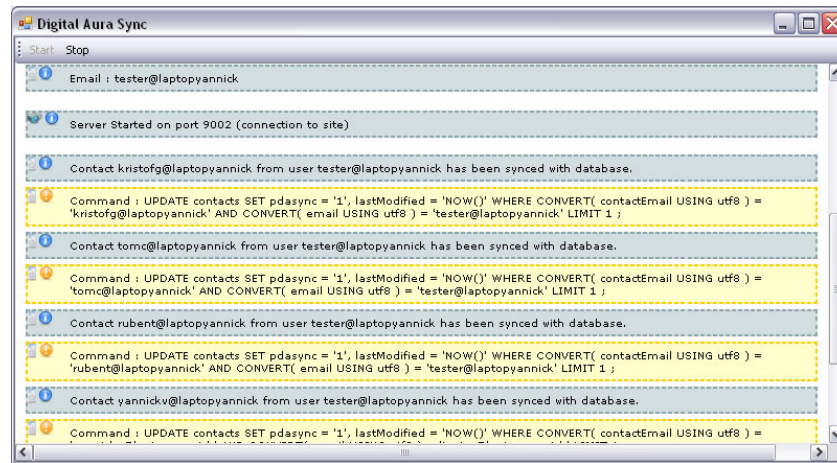
Door op de 'fake event' knop te klikken, bootsen we een aanvraag van de website na. Deze functie is vooral voor debuggingsredenen toegevoegd aan dit venster.

5.3 Synchronisatie Server

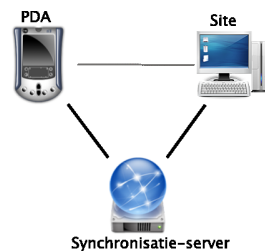
Indien de gebruiker zijn PDA in de cradle¹ steekt en de Digital Aura applicatie opstart, is het mogelijk om de contacten tussen de website en de PDA te synchronizeren. In figuur 5.15 staat een screenshot van de applicatie die de synchronisatie afhandelt. Met behulp van deze server applicatie is het mogelijk om het systeem op fouten te controleren.

De werking van de synchronizatie leggen we uit aan de hand van figuur 5.16. Voor de gebruiker lijkt er een rechtstreekse link te bestaan tussen de PDA en de computer, dit is echter niet het geval. In werkelijkheid worden de contacten van zowel de website als de PDA naar een synchronisatie ser-

¹Dockingstation voor het PDA toestel



Figuur 5.15: Synchronisatie-server voor het synchroniseren van de contacten tussen de site en de PDA applicatie.



Figuur 5.16: Overzicht van de synchronisatie werking.

ver gestuurd. Deze server synchroniseert vervolgens de contacten en stuurt vervolgens de ge-update contactenlijst naar de website en de PDA.

Hoofdstuk 6

Conclusies

In deze thesis hebben we enkele richtlijnen gegeven (zie hoofdstuk 3) die helpen bij het opstellen van een gebruikersprofiel. Voor een User Profile hebben we onderstaande richtlijnen besproken:

- **Gebruiksgemak** : De profielelementen moeten makkelijk en intuïtief aanpasbaar zijn. Bij de afwezigheid van een groot gebruiksgemak zal de gebruiker ervan afzien om veel informatie aan te passen.;
- **Opletten met dubbelzinnige informatie** : Dubbelzinnige profielelementen kunnen de betrouwbaarheid van profile matching sterk ondermijnen. Als verschillende gebruikers een element anders interpreteren kunnen fouten optreden bij de matching.;
- **Standaardwaarden mijden** : Door een standaardwaarde toe te kennen aan een profiel, kan het zijn dat iemand die niets persoonlijks heeft ingegeven, toch gemiddeld scoort bij een profile matching. Dit is een ongewenst neveneffect en moet dus vermeden worden.;
- **Verplichte velden mijden** : Door de gebruiker te verplichten om een lange lijst met profielitems in te vullen, zal deze geënerveerd raken en velden foutief invullen om tijd te besparen. Dit euvel kan voorkomen worden door de gebruiker te laten kiezen welke items hij/zij wenst in te vullen.;

- **Verwelkom (late) wijzigingen** : Het moet mogelijk zijn om profiel informatie aan te passen indien nodig. De kans is bestaande dat een gebruiker na verloop van tijd andere waarden wenst in te vullen voor de profielitems. Als deze mogelijkheid niet bestaat, of als de grens om aan te passen te groot is, zal de gebruiker ervan afzien om het profiel te wijzigen en zal het profiel bijgevolg foute data bevatten.;
- **Maak een verschil tussen exacte en niet-exacte data** : De nominale data moet anders behandeld worden dan de ordinale data;
- **Orden de profielelementen** : De data moeten geordend kunnen worden. Een alfabetische volgorde of een volgorde op datum van aanmaak is vaak niet de beste oplossing. De belangrijkste profiel groepen moeten bovenaan gezet worden terwijl de minder belangrijke data onderaan de weergave gezet moeten worden.;
- **Aangeven van belangrijkheid van profielelementen** : Bepaalde eigenschappen wegen zwaarder door dan andere eigenschappen. De gebruiker moet dit kunnen aangeven zodat een profile matching dichter bij zijn persoonlijke voorkeuren aanleunt. Door ordinale waarden die een belangrijkheid aanduiden te koppelen aan de profielelementen, kan een gebruiker aangeven hoe belangrijk hij/zij dit element vindt.

Tevens hebben we aangetoond dat Context Awareness nuttig is in diverse scenario's. Dit is de reden waarom onze profielen context gevoelig zijn. Deze context gevoeligheid hebben we ingebouwd door de gebruiker te laten kiezen welke informatie hij/zij in een gegeven context wilt vrijgeven. De verschillende contexten bestaan uit *Identity*, *Work*, *Freetime* en *Love*.

Daarna hebben we getracht om profielen met elkaar te vergelijken (zie hoofdstuk 4). Om een matching score te bekomen doorlopen we een kettingstructuur waarbij iedere schakel in de ketting een component voorsteld. Om een score uit de ketting van componenten te bekomen, vermenigvuldigen we iedere component met zijn gewicht in de ketting. Vervolgens sommeren we

alle resultaten van de componenten om een totale uitkomst te bekomen.

Als componenten hebben we voor een Basic Matching, Foaf Matching en Trust Matching gekozen. De Basic Matching component vergelijkt profielementen op basis van de gekozen profielitems behorende tot profielementen. Deze component bepaalt een score door een profielement te vergelijken met een overeenkomstig profielement van een andere gebruiker.

De Foaf Matching houdt rekening met hoe goed gebruikers elkaar kennen. Afhankelijk van de foaf-afstand tussen beide gebruikers geeft deze component een score terug.

Tot slot hebben we de Trust Matching, die het vertrouwen dat gebruikers in elkaar hebben, mee opgenomen in de berekening van een totaal score. Deze component tracht de betrouwbaarheid van een persoon te berekenen door rekening te houden met scores die andere gebruikers aan deze persoon gegeven hebben.

Met deze structuur en componenten behaalden we goede subjectieve resultaten. Bovendien bieden we voldoende flexibiliteit om aanpassingen of uitbreidingen te ondersteunen.

De conclusies die we in deze hoofdstukken gemaakt hebben, hebben we aan de praktijk getoetst gehad. De toetsing aan de praktijk hebben we gedaan door een website en een matching server te implementeren. Via de site kunnen gebruikers een persoonlijk gebruikersprofiel en een zoek-profiel opstellen. Met de server werd er naar interessante gebruikers gezocht om deze resultaten vervolgens op de site aan de gebruiker te presenteren.

De interessante gebruikers worden in een contactenlijst opgeslagen die rechtstreeks naar een PDA overgezet kunnen worden. Op de PDA applicatie staat een Digital Aura applicatie die de gebruikers op de hoogte stelt indien

er een interessant persoon voorbij komt.

User Profile Matching kan echter nog meer uitgebreid worden, in het volgende hoofdstuk doen we enkele voorstellen om de User Profile Matching uit te breiden en te verbeteren.

Hoofdstuk 7

Future Work

Inhoudsopgave

7.1	Intelligent zoekprofiel	77
7.2	Digital Aura op gsm	77
7.3	User Tests	81

Een thesis is een werk dat een jaar tijd in beslag neemt. Het is onmogelijk om alle paden binnen het onderzoeksdomein te bewandelen. Er zullen dus keuzes gemaakt moeten worden over hetgeen onderzocht en geïmplementeerd zal worden. Er zijn echter tal van ideeën die een meerwaarde voor de thesis kunnen bieden, de belangrijkste ideeën lichten we in onderstaande secties toe.

Een eerste uitbreiding op de User Profile Matching is het toevoegen van een *intelligent zoekprofiel*. Door dit *intelligent zoekprofiel* te gebruiken, wordt de gebruiker geholpen bij het updaten en aanpassen van zijn zoekprofiel.

7.1 Intelligent zoekprofiel

Het zoekprofiel wordt door de gebruiker ingevuld. Het systeem kan enkel goed functioneren indien de gebruiker voldoende aandacht besteedt aan het invullen en up-to-date houden van zijn zoekprofiel. Het is echter mogelijk om het systeem te laten helpen in het nemen van beslissingen (*Interface Agents* [44]). Hier kan echter nog verder in worden gegaan door het systeem beslissingen in naam van de gebruiker te laten nemen (*Autonomous Agents* [46]). Deze agenten worden meer in detail besproken in sectie 2.5.

Het implementeren van zo'n functie zal wellicht goed ontvangen worden bij de gebruiker. De aanpassingen die het systeem doorvoert op een zoekprofiel moeten goed aansluiten bij de beslissingen die een gebruiker genomen zou hebben in een bepaalde situatie. Een User Test zoals in voorgaande sectie besproken, is cruciaal daar een agent die foute beslissingen neemt de gebruiker zal storen waardoor deze het systeem niet zal gebruiken. De mogelijkheid om een agent aan of uit te zetten is een must.

Tot nu is de User Profile Matching enkel aanwezig op een website gekoppeld aan een PDA systeem. Deze functionaliteit kunnen we echter uitbreiden naar een gsm systeem.

7.2 Digital Aura op gsm

Om een digital aura te realiseren heeft men nood aan een apparaat dat over voldoende reken Capaciteiten beschikt en tevens geassocieerd kan worden met zijn gebruiker. Indien dit toestel ontdekt word moet men kunnen veronderstellen dat de gebruiker zich in de buurt van dit toestel bevindt.

Op het Expertisecentrum voor Digitale Media aan de UHasselt is een framework voor digital aura ontwikkeld door gebruik te maken van een pda toestel. Een gedetailleerde beschrijving over dit werk is te raadplegen in [49]

en [7]. Vanuit onze website kan er rechtstreeks gecommuniceerd worden met deze applicatie zodat de contactenlijst automatisch overgezet wordt naar de PDA.

Het is echter ook mogelijk om een digital aura te creëren door gebruik te maken van een gsm toestel. Een gsm toestel is typisch een toestel dat de gebruiker dicht in zijn buurt bewaart. Er kan dus verondersteld worden dat de locatie van een gsm geassocieerd kan worden met de locatie van de gebruiker. De reken Capaciteiten van een gsm zijn tegenwoordig ook voldoende om een eenvoudige berekeningen toe te laten. Verder heeft een gsm vaak de beschikking over infrarood of Bluetooth technologie.

In het werk van Nathan Norfleet Eagle wordt met een gsm gewerkt om een digital aura te creëren, hetgeen aantoont dat het mogelijk is om een gsm te gebruiken voor het construeren van een digital aura. Voor de thesis zou het een meerwaarde betekenen moesten we de User Profile Matching kunnen combineren aan een digital aura framework dat met een gsm werkt.

Hier hebben we reeds een basis onderzoek naar vericht waarin we onderzocht hebben hoe we een connectie tussen twee gsm's opzetten en hoe we de User Profile Matching hierin verwerken. De connectie en gegevens uitwisseling gebeurt door middel van Bluetooth technologie. Indien twee gsm's bij elkaar in de buurt komen sturen deze de unieke gebruikersnaam van hun gebruiker door en afhankelijk van het feit of deze gebruiker in de contactenlijst staat, zal de gebruiker bericht worden over de aanwezigheid van een interessante gebruiker.

Verder moeten we ook een afweging maken tussen een server matching die op voorhand gebeurt of het berekenen van een matching score indien er een persoon opgemerkt word door het systeem. We trachten de voor- en nadelen van beide methodes hier met elkaar te vergelijken.

De voor- en nadelen van een gsm matching die gebeurt op het moment zelf heeft als voordeel dat de matching in real-time gebeurt, een persoon die niet in de contacten lijst staat kan toch 'gematcht' worden.

Er zijn echter ook nadelen verbonden aan deze methode. Deze sommen we even kort op :

- De berekeningen en resources zijn redelijk zwaar. Het uitwisselen van een User Profile neemt bandbreedte in en kan dus storend zijn indien de gebruiker gelijktijdig met een andere netwerk activiteit bezig is. Over de berekeningen zijn nog geen details bekend, maar er wordt verwacht dat deze vrij complex zullen zijn en bijgevolg het toestel sterk vertragen.
- Indien er veel gebruikers met een digital aura voorbij komen zal de gsm voortdurend bezig zijn met rekenen en netwerk resources in te nemen terwijl het grootste aantal mensen genegeerd kan worden.

Als we de voor- en nadelen van de server matching bekijken kunnen we stellen dat de nadelen van een gsm matching zijn weggewerkt. We hebben echter wel het nadeel dat de gebruikers in een bestaande contactslist moeten staan. We sommen kort de voordelen van een server matching op :

- De server is veel krachtiger dan een gsm (kan uitgebreid worden met een cluster systeem)
- De gebruiker moet via de site werken, hetgeen een grotere gebruiksvriendelijkheid kan bieden
- Door alle gebruikers te verplichten om geregeld de site te bezoeken kunnen mensen ook via de site met elkaar in contact komen, zonder elkaar hiervoor te moeten tegenkomen.
- Het is makkelijk om 250 000 contacten op te slaan, hier komt men normaal ruim mee toe. Indien dit niet het geval is moet er meer

geheugen beschikbaar zijn (20MB voor 500 000 contacten, 40MB voor 1 000 000 contacten,). Dit behandelen we meer in detail in één van de volgende paragrafen.

Bij server matching doet de server het grootste deel van de matching, dit wil niet zeggen dat de gsm niets meer hoeft te berekenen. Indien we de contacten voorstellen zoals in onderstaande paragrafen besproken, dan staan contacten per context gegroepeerd. Het zou kunnen voorkomen dat een bepaald persoon in een andere context staat met een bijzonder hoge score, er moet dan een afweging worden gemaakt of de gebruiker hiervan op de hoogte moet worden gesteld.

De contactenlijst wordt gegenereerd door de webapplicatie. In een mogelijke implementatie zou de website een xml bestand van de contactenlijst kunnen genereren die op de gsm gebruikt kan worden. Als formaat van het xml bestand stellen we een opdeling voor zoals in onderstaand codefragment in figuur 7.1.

```
<root>
.
.  <group>
.    .  <context>
.    .    .  FreeTime
.    .  </context>
.    .  <persons>
.    .    .  <p id="yannick.vandeweerd" score="87.6" />
.    .    .  <p id="sofie.nolens" score="56.3" />
.    .  </persons>
.  </group>
.  <group>
.    .  <context>
.    .    .  Work
.    .  </context>
.    .  <persons>
.    .    .  <p id="ruben.thys" score="33.2" />
.    .    .  <p id="raf.menten" score="38.9" />
.    .  </persons>
.  </group>
.
</root>
```

Figuur 7.1: Contactenlijst xml structuur.

Indien bovenstaand formaat gebruikt word om contacten op te slaan, kunnen er 250.000 contacten worden opgeslagen in 10MB, hetgeen een aanmerkelijke grootte is. Om een idee te hebben wat 250.000 contacten zijn geven we even enkele cijfers mee die de grote van de huidige social network sites mee.

- *hi5* : 14 000 leden
- *Aslpage* : 25 204 leden
- *Looknmeet* : 176 636 leden
- *Friendster* : 36 miljoen leden of 1 contact op 144 kan worden opgeslagen
- *MySpace* : 106 miljoen of 1 contact op 424 kan worden opgeslagen

Met uitzondering van Friendster en MySpace zouden we iedere gebruiker in onze contactenlijst kunnen opnemen. Het is echter zeer onwaarschijnlijk dat iedere gebruiker als interessant beschouwd wordt, bijgevolg lijkt een contactenlijst van 250.000 contacten ruim voldoende.

Hieruit kunnen we vermoeden dat een server matching aan te raden is in deze situatie, een implementatie met een *user test* zal wellicht meer gedetailleerde en correcte informatie verschaffen.

7.3 User Tests

Het is moeilijk te omschrijven wie een interessant persoon is voor een ander persoon. We kunnen enkel vermoedens hebben over hetgeen een goede matching zou zijn. Een gebruikerstest kan ons helpen bij het zoeken naar hetgeen als interessant beschouwd word door andere gebruikers.

De User Profile Matching die in deze thesis beschreven is helpt de gebruiker om een interessant persoon te vinden. Bij een gebruikerstest moeten er voldoende veel gebruikers aanwezig zijn zodat er voor iedere gebruiker een interessant persoon in het systeem aanwezig is. Hoe meer gebruikers er aanwezig zijn, hoe beter de resultaten zijn. Een gebruikersaantal van 1000 personen zou voldoende groot moeten zijn om het systeem degelijk te testen.

Liefst hebben we gebruikers uit verschillende domeinen. Studenten, zakenmensen, arbeiders en bedienden hebben ieder hun specifieke interesse. Het systeem maakt zich er sterk in door voldoende flexibel te zijn zodat zowel de zakenman een zakenpartner kan vinden en dat de student een nieuwe sport collega leert kennen.

Een mogelijkheid zou zijn om samen met het i-city¹ project te werken. Bij het i-city project zijn voldoende testgebruikers aanwezig om onze applicatie te testen. Daarenboven zijn de gebruikers strategisch uit verschillende domeinen gekozen, hetgeen een vereiste is om onze testing goed te laten verlopen. We verwachten wel van de gebruiker dat ze een basis computerkennis hebben, zodat ze redelijk vlot vertrokken zijn bij het gebruik van de website.

Het testen van het systeem kan best in verschillende stadia verlopen. In een eerste stadium lanceren we de website zodat beta-testers zich vrij kunnen inloggen op het systeem. We laten iedere gebruiker twee weken werken met het systeem. Na deze twee weken vragen we om een questionnaire in te vullen die informeert naar de gebruiksvriendelijkheid en naar “hoe goed de matching werkt”. Verder informeren we of de gevonden personen aan de verwachtingen van de gebruiker voldeden en of ze zonder ons systeem ook contact gemaakt zouden hebben.

In een tweede stadium, nadat de site goed bevonden is, lanceren we tevens de PDA of gsm applicatie. Deze testperiode leggen we vast op drie

¹www.i-city.be

weken. Bij een website verwachten we dat een gebruiker redelijk snel de werking van het systeem zal begrijpen, bij de PDA of gsm zal de leertijd wellicht langer zijn, daarom zal de testperiode ook langer duren. Een tweede reden voor de langere duur van de testperiode is dat gebruikers de eerste week wellicht zeer enthousiast zullen zijn en bijgevolg aangenaam verrast zijn indien er een interessant persoon voorbij komt. We willen echter ook nagaan of het systeem de gebruiker niet gaat storen bij zijn dagelijkse activiteiten zodra het 'nieuwe' ervan af is. Ook deze test sluiten we af met een questionnaire waarin we informeren naar de mening van de gebruiker.

Bij zowel de website als bij de PDA of gsm test zorgen we ervoor dat de gebruikers ons steeds kunnen contacteren via mail of telefoon. Dit zodat we gebruikers kunnen bijsturen indien ze grote fouten maken die de verdere evaluatie kunnen beïnvloeden.

In het begin van de testperiode voorzien we de testgebruikers van een kleine handleiding en een inleidende cursus over het gebruik en doelstelling van het systeem. De ervaren computergebruikers zullen we apart brieven zodat deze de functie van Profile Designer kunnen vervullen.

Bibliografie

- [1] Barnes J. A. Class and committees in a norwegian island parish. Technical report.
- [2] Sebastiano Pizzutilo Addolorata Cavalluzzi, Berardina De Carolis and Giovanni Cozzolongo. Interacting with embodied agents in public environments. Technical report, Department of Informatics, University of Bari, 2004.
- [3] Norman Adams Bill N. Schilit and Roy Want. Context-aware computing applications. *IEEE Workshop on Mobile Computing Systems and Applications*, 1994.
- [4] Rhys Blakely. Valentine hackers target lovelorn surfers. http://business.timesonline.co.uk/tol/business/industry_sectors/technology/article728610.ece, 2006.
- [5] R. et al. Borovoy. Meme tags and community mirrors: Moving from conferences to collaboration. Technical report, 1998.
- [6] Perlman Brehm, Miller and Campbell. Intimate relationships 3e. Technical report, McGraw-Hill Higher Education (New York), 2002.
- [7] Carl Bruninx. Digital aura: Pervasive human-human interaction. Technical report, Universiteit Hasselt, 2006.

-
- [8] Kindratenko V. Cox, D. and D. Pointer. Intellibadge: Towards providing location-aware value-added services at academic conferences. Technical report, 2003.
 - [9] Byrne D. Interpersonal attraction and attitude similarity. *Journal of Abnormal and Social Psychology* 62, pages 713–715, 1961.
 - [10] Byrne D. The attraction paradigm. *Academic Press, New York*, 1971.
 - [11] Dai.sy. Dai.sy, 2007. [Online; accessed 11-April-2007].
 - [12] Libby Miller Dan Brickley. Foaf and the web, 2007. [Online; accessed 18-April-2007].
 - [13] Libby Miller Dan Brickley. Foaf vocabulary specification, 2007. [Online; accessed 18-April-2007].
 - [14] Dan Maynes-Aminzade Daniel Billss, David M. Hilbert. Improving proactive information systems. Technical report.
 - [15] Jackie Neider Tom Davis Dave Shreiner, Mason Woo. *OpenGL Programming Guide*. Addison-Wesley, fifth edition, 2006.
 - [16] Anind K. Dey. Personal and ubiquitous computing. *Springer Volume 5, Number 1*, pages 4–7, 2004.
 - [17] R.I.M. Dunbar. Neocortex size as a constraint on group size in primates. *Journal of Human Evolution* 22.
 - [18] Burgess E. and Wallin P. Homogamy in social characteristics. *American Journal of Sociology* 2, pages 109–124, 1943.
 - [19] Nathan Norfleet Eagle. *Machine Perception and Learning of Complex Social Systems*. PhD thesis, Stanford University, 2005.
 - [20] Pentland A. Eagle N. Reality mining: Sensing complex social systems. Technical report, 2006.

-
- [21] A. Ferscha. Digital aura. Technical report, Institute for Pervasive Computing, University of Linz and Siemens AG.
- [22] H Fisher. Anatomy of love : A natural history of mating, marriage, and why we stray. Technical report, Fawcett-Columbine (New York), 1992.
- [23] Dominik Heckmann. Language technology and cognitive systems. Technical report, Saarland University.
- [24] Yukari Iwatani. Love: Japanese style. Lovegety (<http://www.wired.com/news/culture/0,1284,12899,00.html>).
- [25] Danah Boyd Jeffrey Heer. Vizster: Visualizing online social networks. Technical report, Computer Science Division - University of California - Berkeley, School of Information Management and Systems - University of California - Berkeley.
- [26] join2grow. Join2grow.biz, 2007. [Online; accessed 27-April-2007].
- [27] Nikolaus Kühn. Visualizing large and dynamic communication networks. Technical report, TU Berlin Fakultät IV Institut für Softwaretechnik und Systemgestaltung, 2003.
- [28] Last.fm. Last.fm, 2007. [Online; accessed 11-April-2007].
- [29] Dirk Speelman Katholieke Universiteit Leuven. Inleiding praktische statistiek voor linguïstische toepassingen, 2007. [Online; accessed 18-April-2007].
- [30] David W. McDonald Loren Terveen. Social matching : A framework and research agenda. Technical report, University of Minnesota, University of Washington, 2005.
- [31] Deutsch M. Cooperation and trust. some theoretical notes. Technical report, Nebraska University Press, 1962.

-
- [32] Lorrie Faith Cranor Mark S. Ackerman and Joseph Reagle. Privacy in e-commerce: examining user scenarios and privacy preferences. Technical report, Information and Computer Science in University of California; AT&T Labs-Research, Shannon Laboratory; World Wide Web Consortium, MIT, Cambridge, 1999.
 - [33] et al. McCarthy. Proactive displays and the experience ubicomp project. ubicomp. Technical report, 2003.
 - [34] Ziegler C. N. and Golbeck J. Investigating correlations of trust and interest similarity. 2006.
 - [35] Jon Orwant. Heterogeneous learning in the doppelgänger user modeling system. *User Modeling and User-Adapted Interaction 4*, pages 107–130, 1995.
 - [36] Goodman E Paulos E. The familiar stranger : Anxiety, comfort and play in public spaces. Technical report, 2004.
 - [37] piccolo. Piccolo toolkit, 2007. [Online; accessed 20-Mai-2007].
 - [38] ratepoint. ratepoint, 2007. [Online; accessed 11-April-2007].
 - [39] Marsh S. *Formalising Trust as a Computational Concept*. PhD thesis, University of Stirling, Department of Mathematics and Computer Science, 1994.
 - [40] Milgram S. The individual in a social world : Essays and experiments. Technical report, 1977.
 - [41] Bettina Berendt Sarah Spiekermann, Jens Grossklags. E-privacy in 2nd generation e-commerce: Privacy preferences versus actual behavior. Technical report, Humboldt University Berlin Institute of Information Systems, 2001.
 - [42] Robbie Schaefer. Fuzzy evaluation of user profiles. Technical report.

-
- [43] Albrecht Schmidt. Implicit human computer interaction through context. Technical report, 2000.
 - [44] Berardina De Carolis Sebastiano Pizzutilo and Fiorella de Rosis. Cooperative interface agents. Technical report, Department of Computer Science - University of Bari, Italy.
 - [45] Peter Braun Steffen Kern and Wilhelm Rossak. Mobisoft : An agent-based middleware for social-mobile applications. Technical report, University Jena, Computer Science Department, The agent factory GmbH, 2006.
 - [46] Henry Lieberman Sybil Shaerin. Intelligent profiling by example. Technical report.
 - [47] Newcomb T. The acquaintance process. 1961.
 - [48] Andrew T.Fiore and Judith S. Donath. Online personals : An overview. Technical report, MIT MEdia Lab, 2004.
 - [49] Kris Luyten Tim Clerckx, Geert Houben and Karin Coninx. Extending social networks with implicit human-human interaction. Technical report, Hasselt University, EDM, IBBT and transnationale Universiteit Limburg.
 - [50] Nils Behrens Tom Nicolai, Eiko Yoneki and Holger Kenn. Exploring social context with the wireless rope. Technical report, TZI Wearable Compting Lab, Universität Bremen, University of Cambridge.
 - [51] Wikipedia. Dunbar’s number, 2007. [Online; accessed 9-March-2007].
 - [52] Wikipedia. Foaf (software) — wikipedia, the free encyclopedia, 2007. [Online; accessed 21-April-2007].
 - [53] Wikipedia. Small world experiment — wikipedia, the free encyclopedia, 2007. [Online; accessed 17-May-2007].

- [54] Wikipedia. Social network — wikipedia, the free encyclopedia, 2007.
[Online; accessed 9-March-2007].
- [55] Ka-Ping Yee. Animated exploration of dynamic graphs with radial layout. Technical report, University of California, Berkeley, 2001.

Auteursrechterlijke overeenkomst

Opdat de Universiteit Hasselt uw eindverhandeling wereldwijd kan reproduceren, vertalen en distribueren is uw akkoord voor deze overeenkomst noodzakelijk. Gelieve de tijd te nemen om deze overeenkomst door te nemen, de gevraagde informatie in te vullen (en de overeenkomst te ondertekenen en af te geven).

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:

User Profile Matching

Richting: **Master in de informatica**

Jaar: **2007**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Ik ga akkoord,

Yannick Van de Weerd

Datum: **23.05.2007**