



Universiteit Utrecht

De bijdrage van *mobile device* tijd-locatie metingen voor sociaal wetenschappelijke netwerk onderzoek

Exploratief onderzoek naar een nieuwe methode om onderzoek te doen naar sociale netwerken. Binnen dit onderzoek is gebruik gemaakt van de data afkomstig uit de Tabiverplaatsing applicatie. Het gebruik van deze data is mede te danken aan een samenwerking tussen de Universiteit Utrecht en het Centraal Bureau voor de Statistiek.

Studenten

Hochun Wu 4278267

Gerrit Wijnands 5732549

Begeleiders:

Dr. Peter Lugtig

Prof. dr. Barry Schouten

29 juni 2018

Inhoudsopgave

<i>Introductie</i>	3
Mobiliteit en netwerk	4
Ontwikkeling van ruimtelijke netwerkanalyse	5
Ontwikkeling van netwerkonderzoek	6
<i>Achtergrond</i>	7
Persoonlijke netwerken organiseren zich rondom gezamenlijke activiteiten	7
Tijdsbesteding	8
Verplaatsing, mobiliteitspatronen en vriendschap	9
Omvang van het netwerk	11
Tabi-verplaatsing applicatie van het Centraal Bureau voor de Statistiek	11
Mobiliteit als indicatie van netwerk karakteristieken	13
<i>Methodologie</i>	14
Alleen kwantitatieve metingen	15
Nieuw wetenschappelijk inzicht op netwerken	16
Hoe meet de applicatie?	17
Visuele informatie binnen de kaartjes	19
<i>Tekortkomingen methodologie en data van Tabi-verplaatsingen</i>	20
Foutieve patronen: meetfouten	20
Foutieve patronen: oorzaken	23
<i>Resultaten</i>	26
Analyse van de mobiliteitsdata	26
<i>Conclusie & discussie</i>	33
<i>Referenties</i>	36

Introductie

In de wetenschappelijke wereld is er al veel onderzoek gedaan naar sociale netwerken. Een definitie die kan worden gegeven aan een sociaal netwerk is: ‘Een groepering van mensen met wie een persoon (de centrale actor) min of meer duurzame banden onderhoudt voor de vervulling van noodzakelijke levensbehoeften’ (Hendrix, 2001). Onder andere door de snelle ontwikkeling van sociale media zijn er verschillende perspectieven ontwikkeld om analyse te doen naar een individueel netwerk. Zo is het door vernieuwde technologie mogelijk een virtueel netwerk te creëren met mensen die een actor nog nooit gezien heeft. Buiten deze online technieken zijn er ook nieuwe technieken welke het mogelijk maken om ruimtelijk onderzoek te doen naar persoonlijke netwerken.

Netwerkonderzoek richt zich voornamelijk op concepten zoals sociaal kapitaal (= ‘grondstof’ van de sociale relatie). Doordat persoonlijk kapitaal mede afhankelijk is van iemands sociale netwerk, zijn de persoonlijke sociale netwerken steeds belangrijker geworden. Sociale relaties kunnen als een middel worden ingezet om een bepaald doel te behalen (bijvoorbeeld: vinden van een baan). Ook bieden sociale netwerken verklaring voor individuele gedragingen en attitudes. Kortom, vanuit sociologisch perspectief worden sociale netwerken als sterke determinanten beschouwd voor individuele fenomenen. Dit maakt het relevant om te onderzoeken of met een ruimtelijke/positionele vorm van data (GPS-data) ook uitspraak gedaan kan worden over sociale netwerk karakteristieken op individueel niveau.

Sociale netwerken kennen diverse aspecten zoals omvang en diversiteit. Ze worden gebruikelijk met drie methodes gemeten. Een eerste methode is het gebruik van een vragenlijst waarin de hoeveelheid vrienden en de intensiteit kan worden gemeten. Dit is vooral een methode die in het begin van ruimtelijk netwerkonderzoek wordt toegepast. Een tweede methode is het gebruik van online data. Deze recente methode gaat vooral in online netwerken door bijvoorbeeld het aantal Facebook of LinkedIn connecties te gebruiken ter indicatie van sociale netwerken. Een derde, nog vrij nieuwe methode, is met het gebruik van locatie-gebonden data. Tegenwoordig zijn applicaties vaak locatie-gebonden (bijvoorbeeld Facebook, Twitter en Instagram). Dit betekent dat de applicatie jouw locatie bekend maakt door het gebruik van sensoren binnen je *smartphone*. Onderzoekers kunnen deze data gebruiken om te kijken naar welke verschillende locaties een individu geweest is (Eagle et al. 2009). Dit is voor bedrijven een belangrijke vorm van data, omdat op deze manier goed

ingespeeld kan worden op consumenten. Echter voor netwerk onderzoek is locatie-gebonden data binnen de wetenschappelijke literatuur nog weinig toegepast.

Methodes van onderzoek naar sociale netwerken zijn door de laatste 20 jaar sterk ontwikkeld. Terwijl vroeger respondenten werd gevraagd hun verplaatsingen schriftelijk bij te houden in de vorm van een dagboek, is het tegenwoordig mogelijk om met behulp van applicaties verplaatsingen te meten aan de hand van GPS-locaties die bijgehouden worden door een smartphone. Deze relatief nieuwe methode is waar dit onderzoek zich bij aansluit. Zo zal er in dit onderzoek gebruik gemaakt worden van GPS-data afkomstig uit de ''Tabi-verplaatsingen applicatie'' die ontworpen is door het Centraal Bureau voor de Statistiek (=CBS). Deze applicatie is een nieuw ontworpen meetinstrument welke mensen kunnen downloaden op hun *smartphone*.

Mobiliteit en netwerk

Sociale netwerken en mobiliteit zijn nauw aan elkaar gerelateerd. Zo zorgen sociale relaties ervoor dat je je gaat verplaatsen. Verder veroorzaken netwerken voor vaste bewegingspatronen, doordat je regelmatig contact houdt met vrienden en familie en je telkens verplaatst naar de locatie waar je vrienden zijn (Belot & Ermisch, 2009). Ook vaste routines zoals werk of school creëren een vaste bewegingspatroon. Hierbij is wel de fysieke beweging een belangrijke factor. Zo kan gesteld worden dat deze relatie tussen sociale netwerken en mobiliteit mutueel lijkt; zonder mobiliteit is het niet mogelijk een sociaal netwerk op te bouwen en vice versa. Daarnaast is het zo dat hoe verder je je verplaatst, hoe minder waarschijnlijk het is dat je connecties van binnen jouw sociale netwerk tegenkomt (Musolesi, 2007). Je sociale netwerk zal groter zijn rondom je woon- en werkplaats dan in een andere provincie of in het buitenland. Kortom, verplaatsingen in de vorm van mobiliteit is een invloedrijke factor voor de vorm en diversiteit van een netwerk. Dat het netwerk invloed heeft op ons bewegingsgedrag blijkt uit onderzoek (Cho et al, 2011). Uit dit onderzoek komt naar voren dat 10 tot 30 procent van ons bewegingsgedrag verklaard kan worden aan de hand van sociale relaties die je hebt, voor periodiek gedrag verklaard het netwerk maar liefst 50 tot 70 procent (Cho, Myers en Leskovec, 2011). Sociale netwerken blijken dus een invloedrijke factor te zijn wat betreft het bewegingsgedrag van mensen. GPS-data is hierbij een goede vorm van data om fysieke beweging in kaart te brengen.

Ontwikkeling van ruimtelijke netwerkanalyse

Zoals eerder vermeld zijn er voor netwerkanalyse verschillende methodes ontwikkeld. Echter kennen deze methodes wel enkele gebreken. Bijvoorbeeld het vragen van een respondent om al zijn activiteiten bij te houden in een dagboek over een langere periode is niet realistisch. Hierom kunnen respondenten minder welwillend worden om mee te doen aan dergelijke onderzoeken. Daarbij lijken respondenten bepaalde (triviale) verplaatsingen, zoals het uitlaten van een hond, niet altijd te noteren.

Door technologische ontwikkelingen is het tegenwoordig mogelijk om via GPS-signalen van een *smartphone* verplaatsingen bij te houden. Op deze wijze hoeft een respondent zelf geen dagboek meer bij de hand te houden. Deze geautomatiseerde methode zal waarschijnlijk meer valide resultaten opleveren, zo blijkt dat tijdens het individueel registreren van verplaatsingen er een duidelijk verschil is binnen de kwaliteit van de data van automatische metingen en handmatige metingen (Hoogendoorn-Lanser, Schaap, Kalter & Coffeng, 2013). Daarbij vermindert geautomatiseerd onderzoek, buiten de meetfouten, ook de kosten van het onderzoek.

De Tabi-verplaatsingen applicatie is een medium die verplaatsingen digitaal bij kan houden. Hiermee wordt het mogelijk om op basis van het reisgedrag te zien welk vervoermiddel het meest wordt gebruikt, welke wegen vaak worden gebruikt en waar de woon- en werkplek zijn van de respondent. De data welke afkomstig zijn uit de Tabi-verplaatsing applicatie kunnen zo nuttig zijn voor derde partijen binnen de maatschappij. Deze mobiliteitsdata kunnen bijvoorbeeld basis bieden voor het Ministerie van Rijkswaterstaat om een gepast beleid te ontwikkelen wat in de toekomst file problemen zou moeten voorkomen. Met de data wordt immers helder waarom respondenten zich op de desbetreffende snelweg bevinden.

Verder is genoemd dat momenteel vooral onderzoek gedaan naar online sociale netwerken, dit komt door de opkomst van grote social media forums zoals Facebook en LinkedIn. Deze onderzoeken zijn zeker van belang voor de kennis omtrent netwerken, maar ook deze methode kent zijn nadelen. Zo heerst er een algemeen discussiepunt wanneer het gaat om fysieke beweging. Binnen het opstellen van online netwerken bestaat de kans dat een persoon zijn netwerk verkeerd wordt neergezet of dat er zogezegd "gaten" in het netwerk zitten. Het ware netwerk van de respondent kan hierdoor niet uit de data worden afgeleid. Dit komt onder andere doordat een individu ook relaties kan hebben met mensen die geen gebruik maken van social media (Valenzuela, Park & Kee, 2009). Om een totaalbeeld te krijgen van de karakteristieken (diversiteit en vorm) van een netwerk is het belangrijk fysieke

beweging in de vorm van mobiliteit mee te wegen binnen een analyse. Buiten dat het online netwerk van een persoon veelzeggend is over bijvoorbeeld persoonlijk sociaal kapitaal (Ellison, Steinfield & Lampe, 2007), biedt mobiliteit nog nieuwe inzichten binnen het daadwerkelijke netwerk van een persoon. Ook laat het zien in hoeverre een persoon op fysieke wijze connecties kan maken, wat met online netwerkanalyse niet mogelijk is. Een nadeel van GPS-data is dat er een selectie-effect optreedt. Dit treedt op omdat je alleen mensen aantrekt die in het bezit zijn van een smartphone en die ook goed weten te gebruiken. Je laat daarbij een groep die geen smartphones bezit of gebruikt buiten beschouwing.

Een zwakte van online netwerkanalyse is dat je een deel van sociale netwerken niet meet, namelijk alle contacten die je niet online hebt (= offline netwerk). GPS-data zou de tekortkomingen van online netwerkanalyse kunnen verhelpen. GPS-data is nuttig om de offline netwerken van mensen te meten omdat het bij deze data gaat om fysieke verplaatsingen en ook face-to-face contact, wat aan online netwerkanalyses ontbreekt. Deze alternatieve vorm van data zou daarom een belangrijke bijdrage leveren aan sociaal netwerk onderzoek.

Ontwikkeling van netwerkonderzoek

Binnen deze thesis wordt onderzocht hoe concreet er onderzoek gedaan kan worden naar ruimtelijke netwerken, met behulp van GPS-data. Belangrijk bij dit onderzoek is dat er kritisch naar het functioneren van de, door het CBS ontwikkelde, Tabi-verplaatsing applicatie wordt gekeken. Om effectief onderzoek te doen naar sociale netwerken is het noodzakelijk dat er bepaalde randvoorwaarden worden gesteld aan de data die geproduceerd wordt door de Tabi-verplaatsingen applicatie. Zo is het belangrijk dat de respondent altijd de applicatie aan heeft staan en zijn smartphone mee heeft. Daarnaast is het van belang dat er geen tot weinig meetfouten en missende data geproduceerd worden. Als aan deze voorwaarden wordt voldaan betekent het dat de Tabi-verplaatsingen applicatie kwalitatieve data produceert dat gebruikt kan worden om sociale netwerken te bestuderen. In ons onderzoek zal er gekeken worden naar of de applicatie kan voldoen aan de bovengenoemde randvoorwaarden en hoe er aan die voorwaarden kunnen worden voldaan.

Uiteindelijk is het doel van dit onderzoek om inzicht te krijgen in de waarde van GPS-data voor sociaal netwerk onderzoek. Geeft deze vorm van data genoeg relevante informatie waaruit, op basis van literatuur, netwerk karakteristieken kunnen worden afgeleid. Zo gaat dit onderzoek verder in op ontwikkeling van ruimtelijk netwerkonderzoek, waarbij de factor mobiliteit de belangrijkste is.

Achtergrond

Ondanks dat onderzoek met GPS-data nog redelijk nieuw is, is er al wel literatuur bekend over sociaal netwerkonderzoek, mobiliteit en de karakteristieken van sociale netwerken. In de komende paragraaf wordt er dieper ingegaan op de literatuur over mobiliteit en sociale netwerken. Vervolgens wordt er ook dieper ingegaan op de Tabi-verplaatsingen applicatie.

Persoonlijke netwerken organiseren zich rondom gezamenlijke activiteiten

De mate van diversiteit van een netwerk is afhankelijk van de karakteristieken van de relaties van een persoon. Over het algemeen geldt dat mensen voornamelijk de voorkeur geven aan connecties met wie zij verschillende persoonlijke kenmerken delen (Tajfel & Turner, 2004). Hierbij kan gedacht worden aan gelijke status, maar ook aan gelijke demografische kenmerken zoals ras, religie of sekse. Vanuit dit idee kan gesteld worden dat het netwerk van mensen over het algemeen een vrij lage mate heeft van diversiteit, waarmee bedoeld wordt dat de connecties binnen zijn/haar netwerk weinig verschillen wat betreft persoonlijke kenmerken. Echter, Tajfel en Turner beschrijven de mens dan ook vooral als een volledig rationeel wezen binnen hun homogeniteitstheorie (Tajfel & Turner, 2004).

Deze persoonlijke voorkeuren ontwikkelen zich niet volledig onafhankelijk van externe contextuele invloed. Deze invloed komt terug binnen het idee van de focus-theorie van Feld. Binnen zijn onderzoek *The focused organization of social ties* wordt uitgelegd hoe een deel van het menselijke netwerk tot stand komt. Binnen zijn idee zijn er terugkerende *Foci*, wat men kan beschouwen als fysieke of sociale omgeving waarbinnen gemeenschappelijke en gezamenlijke activiteiten plaatsvinden (Feld, 1981). *Foci* kunnen bijvoorbeeld scholen, werkplekken of sportverenigingen zijn. Feld gaat ook uit van persoonlijke voorkeuren voor overeenkomsten, maar de connecties binnen persoonlijke netwerken vinden onder andere plaats door het delen van een gezamenlijke activiteit(en) (= *focus/foci*). Doordat verschillende mensen dezelfde *foci* delen (zoals teamgenoten of collega's) ontstaat er een connectie tussen deze personen. De mate van intensiteit van deze connectie hangt voor een groot deel af van de gependeerde tijd binnen de gemeenschappelijke *foci*, de hoeveelheid mensen die binnen de *foci* aanwezig zijn en de mate waarin een *foci* toegankelijk is. Een student op de universiteit zal zo niet met iedereen een connectie hebben doordat er heel veel studenten op de universiteit zijn, doordat niet alle studenten op hetzelfde moment aanwezig kunnen en zullen zijn en studenten op de universiteit toegang hebben tot een deel van de aangeboden cursussen.

De *foci* zoals hierboven beschreven, kunnen in het kader van dit onderzoek geïnterpreteerd worden als *stop* (= respondent bevindt zich langere tijd op eenzelfde punt). De plekken waar respondenten vaker komen zullen waarschijnlijk plekken als werk, thuis of vrijetijdsactiviteiten zijn. Doordat het binnen de applicatie mogelijk is het doel van de *stop* te beschrijven kan worden opgemaakt of de *stop* een gezamenlijke activiteit is. Deze *stops* zijn weer af te leiden uit het mobiliteitspatroon en de sensor-data van een persoon. Op basis van de frequentie van gezamenlijke activiteit(en) kan weer uitspraak gedaan worden over de diversiteit van een netwerk. Door het homogeniteitsprincipe en het feit dat mensen met gemeenschappelijke *foci* grote kans hebben elkaar te kennen kan uitspraak gedaan worden over de mate van netwerkdiversiteit van een persoon. Deze aanname wordt empirisch ondersteund door een recentelijke onderzoek met behulp van Facebook. Op dit sociale medium is het mogelijk om te delen met connecties op wat voor locatie je bent (bijvoorbeeld een restaurant of een musical). Gao, Tang en Liu (2012) hebben deze verschillende gedeelde locaties van connecties onderling geanalyseerd en de analyse bevestigde dat actoren die in connectie met elkaar staan vaker op dezelfde plekken komen dan mensen die niet in connectie zijn. Ook wanneer realistische modellen worden gesimuleerd blijkt dat vriendschappen/netwerk connecties zich blijven ontwikkelen door toedoen van gemeenschappelijke activiteiten/locaties (Musolesi & Mascolo, 2007).

Tijdsbesteding

Binnen de *Focus-theory* van Feld is de mate van tijdsbesteding een voorspeller van de hoeveelheid connecties binnen deze focus. Om de diversiteit en vormgeving van het netwerk van een individu inzichtelijk te maken is het ook belangrijk te weten hoeveel tijd zij in welke activiteit steken. Wanneer bekend is in wat voor activiteiten individuen participeren en hoeveel tijd zij hieraan spenderen, is het mogelijk om op basis hiervan de intensiteit te weten te komen van een individu. De mate van intensiteit van tijdsbesteding is van invloed op hoe het netwerk van een individu vorm krijgt (Passy & Giugni, 2001). Gedurende de activiteiten waarin een individu participeert, houdt een individu in zijn gedragingen rekening met de invloed van gezag (autoriteit), de collectieve afweging (bijvoorbeeld: een staking) en individuele voorkeuren. Dit zijn vooral individueel afgewogen keuzes. Echter, een individu wordt vaak buiten de individuele afwegingen beïnvloed door de actoren die zich in het netwerk van de individu bevinden. Bandura (1986) introduceert hierbij ook de sociale cognitieve theorie, welke voorspelt dat er meerdere blootstellingen van bepaald gedrag moeten plaatsvinden voordat een individu dit gedrag gaat overnemen. Wanneer uit de tijdsbesteding

analyse van een respondent blijkt dat deze veel tijd spendeert aan zijn sportvereniging, kan ook gesteld worden dat deze veel connecties heeft binnen deze vereniging. Op deze manier geeft de tijdsbesteding van een netwerk inzicht over de diversiteit en grootte van een netwerk.

Door bovenstaande afleiding is het ook mogelijk betekenis te geven aan locaties die binnen de mobiliteitspatronen in de data aangetroffen worden. In tijdsbesteding studies komt namelijk naar voren hoe mensen hun dag indelen en waar zij hun tijd aan besteden. Er wordt aan de participanten gevraagd bij te houden wat ze elke dag precies doen. Deze methodologische werkwijze komt overeen met de dataverzameling van de Tabi-verplaatsing applicatie, hier zal later nog op teruggekomen worden.

Tijdsbesteding studies zijn voor onze studie erg bruikbaar omdat deze te verwachten patronen kunnen aangeven die wij in onze dataset aantreffen (bijvoorbeeld naar werk, school of supermarkt). In een onderzoek van Hamermesh et al. (2005), waarbij Amerikaanse huishoudens hun dagelijkse tijdsbesteding hebben gehouden, is dit geïllustreerd. Zo blijkt dat verreweg de meeste tijd werd besteed aan persoonlijke zorg, ruim 9 uur. Dit komt voornamelijk omdat slapen onder deze categorie valt. Verder wordt er de meeste tijd besteed aan werk gerelateerde activiteiten. Bij mannen is dit gemiddeld 4 uur per dag en bij vrouwen 2.5 uur. Dit uren-aantal voor vrouwen is lager dan verwacht, omdat de data in 1997 en 2002 zijn verzameld. Ook kan meespelen dat in deze tijd de traditionele genderrollen nog sterker heersten dan tegenwoordig (Bianchi, 2000). Ook moet er rekening worden gehouden met deelname van telkens verschillende leeftijdsgroepen, dus ook mensen zonder een baan. Ten slotte wordt er gemiddeld 4 uur besteed aan vrijetijdsactiviteiten. Data van het *Bureau of Labor Statistics* in Amerika ondersteunen het genoemde onderzoek naar tijdsbesteding. Zo zijn uit deze gegevens af te leiden dat persoonlijke verzorging, werk en vrijetijdsbesteding het meest voorkomend zijn. Aan persoonlijke verzorging wordt gemiddeld 9 uur besteed, aan werk gemiddeld 6.5 uur en aan vrije tijd ruim 4 uur (BLS, 2016).

Aan de hand van tijdsbesteding wordt zo duidelijk wat mensen ongeveer gedaan hebben op een dag. Daarbij krijgt men ook inzicht in de mobiliteitspatronen van respondenten. Tijdsbesteding kan bij netwerkanalyse dus een belangrijke factor van invloed zijn, omdat het mogelijke inzichten geeft van de mobiliteit van een persoon.

Verplaatsing, mobiliteitspatronen en vriendschap

Vaak wordt er gedacht dat onze verplaatsingen en mobiliteitspatronen veel vrijheden en variatie bevatten, dit blijkt in de realiteit toch anders te zijn. Individuen worden in hun verplaatsingen beperkt door geografische en sociale beperkingen. Te verwachten valt dat

individuen sterke periodieke gedragingen vertonen in hun verplaatsingen omdat ze heen en weer reizen van werk naar huis (Eagle et al., 2009). Verplaatsingen worden beperkt op geografische wijze omdat we op één dag maar een bepaalde afstand kunnen afleggen (Gonzalez et al., 2008). Uit onderzoek blijkt dat het veel waarschijnlijker is dat je vrienden hebt die geografisch dichterbij leven dan ver weg. Hoe ons sociaal netwerk eruit ziet wordt beïnvloed door onze geografische positie (Liben-Nowell et al., 2005).

Mobiliteitspatronen van mensen zijn tijd en locatie gebonden. Dit betekent dat rond bepaalde tijden vaak dezelfde locaties worden gezocht en dat men frequent terugkeert naar deze locaties. Ondanks de diversiteit van bezochte locaties van individuen, volgen zij vaak simpele herhaalbare mobiliteitspatronen (Gonzalez et al., 2008).

Daarnaast wordt mobiliteit ook gevormd door ons sociale netwerk, individuen zijn eerder geneigd om locaties te bezoeken die vrienden of mensen die op ons lijken voorheen hebben bezocht (Goldenberg et al., 2009).

In een onderzoek van Stanford university naar mobiliteit en vriendschap is er gebruik gemaakt van location-based social networks data (Cho et al., 2011). Deze data komen voort uit applicaties die de locatie van de gebruiker meeneemt (bijvoorbeeld: Instagram, Twitter, Facebook). Bij mobiele GPS-data kan het netwerk van een individu ook afgeleid worden aan de hand van locaties, bij *location-based* data is het sociale netwerk expliciet aangegeven. Opvallend in het onderzoek is dat de afstand geen invloed heeft op de kans dat je een vriend gaat bezoeken, terwijl we een negatief verband verwachten tussen de twee variabelen (Hoe verder je reist, hoe minder vrienden je hebt in de omgeving). Hun onderzoek toont aan dat invloed op de mobiliteit van een individu relatief toeneemt naarmate zijn/haar vriend verder weg woont. Uiteindelijk komen in het onderzoek twee modellen naar voren: *Periodic Mobility Model* (=PMM) en *Periodic & Social Mobility Model* (=PSMM). Het PMM gaat van het idee uit dat verplaatsingen vooral gebaseerd is op een kleine set locaties. Onder deze set van locaties verstaan we ‘werk’ en ‘thuis’ en afhankelijk van de tijd op de dag zal de verplaatsingen van een persoon gecentreerd zijn op basis van één van de twee locaties. Het PSMM model is een uitbreiding van het PMM model, maar in dit model wordt ook het sociale netwerk van het individu meegenomen door middel van *check-in* locaties. Dit zijn speciale locaties van applicaties die meten waar een persoon is geweest. Daarnaast tonen zij in hun onderzoek aan dat 10 tot 30 procent van onze dagelijkse beweging verklaard kan worden op basis van onze sociale netwerk. Maar liefst 50 tot 70 procent van onze mobiliteit wordt verklaard aan de hand van periodieke bewegingsgedrag (Cho et al., 2011). Dit betekent dat grotendeels van onze mobiliteit bestaat uit korte en periodieke bewegingspatronen.

Daarnaast zou dit betekenen dat bewegingspatronen vaak kort en rondom onze woonplek is en er relatief weinig lange afstanden worden afgelegd.

Omvang van het netwerk

Bovenstaande paragrafen maken duidelijk dat netwerkdiversiteit en mobiliteit vanuit verschillende perspectieven in verband staan met elkaar. Buiten de diversiteit van een netwerk kan aan de hand van mobiliteitspatronen ook analyse gedaan worden naar de omvang van een netwerk. Met netwerkomvang wordt het aantal actoren binnen persoonlijke netwerken bedoeld.

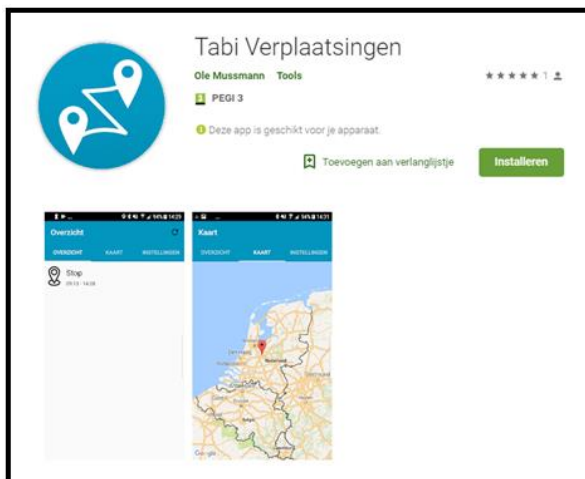
De mate van participatie hangt ook sterk af van de urbanisatiegraad van het leefgebied van een persoon (Tijhuis et al., 1992). Mensen binnen rurale gebieden, of mensen welke meer lokaal dicht bij hun woning leven, kunnen hierdoor een kleinere omvang hebben van hun netwerk. Echter moet er wel sceptisch naar deze veronderstelling gekeken blijven worden, omdat netwerk niet causaal verklaard hoeft te worden door de urbanisatiegraad van een gebied. Mensen breiden ook zelf hun netwerk uit. Hiervan is het van belang dat zij de structuur van een netwerk snappen, zij weten hoe ze zich binnen dit netwerk moeten gedragen volgens de sociale normen en zij de link tussen zichzelf en de structuur van het netwerk weten te behouden (Jack & Anderson, 2002). Volgens Jack en Anderson (2002) maakt urbanisatiegraad hierbij weinig uit. Zij hebben in hun onderzoek voornamelijk de mens als individu geanalyseerd en hebben daaruit de hiervoor genoemde voorwaarden ontwikkeld.

Tabi-verplaatsing applicatie van het Centraal Bureau voor de Statistiek

Het Kennisinstituut voor Mobiliteitsbeleid (=KIM) levert ieder jaar nieuwe inzichten en gegevens aan bij het ministerie van Infrastructuur en Waterstaat (=IenW). Hiervoor wordt onder andere onderzoek gedaan naar de mobiliteit van goederen-, personen- en openbaarvervoer (KIM, 2018). Om kennis en relevante ontwikkelingen openbaar te maken bij IenW, publiceren zij ieder jaar een mobiliteitspanel waarin gegevens van reisgedrag en mobiliteitspatronen. Deze gegevens zijn belangrijk, omdat deze van invloed zijn op het beleid dat de overheid gaat stellen. Buiten het KIM geeft IenW ook aan andere instanties opdracht om bepaald onderzoek te doen. Zo is het IenW op dit moment geïnteresseerd in de individuele mobiliteit van Nederlandse burgers. Om hier duidelijkheid over te krijgen heeft IenW beroep gedaan op het CBS. Zo is de taak voor CBS om de mobiliteitspatronen van individuen in kaart te brengen.

Het CBS doet dit door het ontwikkelen van een eigen mobiele applicatie die via GPS-sensoren bijhoudt waar een persoon op een dag allemaal geweest bent. Er is voor gekozen om een eigen applicatie te ontwikkelen, vanwege het feit dat dit veel flexibiliteit biedt in de methodologie van de applicatie. Hiermee wordt bedoeld dat de ontwerpers en ontwikkelaars zelf de keuzes binnen de software voor de applicatie kunnen maken en dit heeft als voordeel dat zij hierbij niet afhankelijk zijn van een andere partij. Ook kan door het ontwikkelen van de applicatie bij fouten of onwenselijk specificaties in de applicatie ingegrepen worden. Dit onderzoek kunnen zij doen, omdat zij ook alles zelf geprogrammeerd hebben en dus inzicht hebben in de vormgeving van de software. Het grote voordeel van het zelf ontwikkelen van een applicatie is het inzicht dat de ontwikkelaars van CBS zelf hebben in de applicatie. Daarbij zijn andere, al bestaande, GPS-data applicaties erg hoog in kosten om te gebruiken. Alles wat binnen de software gebeurt en specificaties oplevert hebben de ontwikkelaars zelf in de hand. Deze applicatie is openbaar voor elke eigenaar van een *smartphone*.

Figuur 1: Tabi-verplaatsing applicatie in de Google Play Store



Op dit moment zijn er verschillende *testrondes* geweest met de applicatie. Met deze *rondes* is gecontroleerd of de applicatie de mobiliteit goed bijhoudt. Het algemene beeld na deze testen is dat er nog enkele meetfouten binnen de applicatie plaatsvinden. Een voorbeeld hiervan is een GPS-coördinaat dat aangeeft dat een respondent midden op zee is, zonder dat hier omliggende coördinaten rondom aanwezig zijn. Een sensor-datapunt met zulke kenmerken lijkt erg onwaarschijnlijk.

Eerder is aangegeven dat dit bacheloronderzoek erop gericht is onderzoek te doen naar de ruimtelijke netwerken met behulp van GPS-data. Ook de methodologische fouten

binnen de Tabi-verplaatsing applicatie zijn hierbij van belang. Zo zal er binnen dit onderzoek gezocht worden naar een toepasbare benadering voor deze meetfouten binnen de data. Door uitgebreid de stappen binnen het selectieproces van de data te beschrijven, wordt niet alleen de kans op valide en betrouwbare data vergroot, maar ook de repliceerbaarheid van dit onderzoek. Hierbij moet vermeld worden dat de applicatie in de afgelopen maanden door is ontwikkeld en dat sommige meetfouten wellicht niet meer voorkomen binnen de data. Door de focus op de meetfouten houdt deze thesis zich meer bezig met de analyse van de GPS-data die afkomstig zijn uit de applicatie, dan met de software instellingen binnen de applicatie.

Mobiliteit als indicatie van netwerk karakteristieken

Om te onderzoeken hoe sociale netwerken afhankelijk kunnen zijn van de mobiliteit van mensen, richt dit onderzoek zich op hoe GPS-data gebruikt kan worden om sociale netwerken te bestuderen. Om de koppeling tussen de ruimtelijke beweging en het sociale netwerk wordt vooral gelet op de *stop* van een respondent. De context rondom de zogenoemde *stops* kan namelijk gezien worden als voorspeller van de vormgeving en diversiteit van een netwerk.

Op basis van deze *stops* kunnen meerdere aspecten van de ruimtelijke beweging onderzocht worden. Vanuit deze aspecten zijn verschillende indicatoren afgeleid, welke allemaal onafhankelijk informatie bieden over de structuur van het ruimtelijke netwerk. De mogelijke indicatoren voor de netwerkanalyse zijn te vinden in **tabel 1**. Hoe deze een verdere rol krijgen binnen deze thesis wordt uitgebreid beschreven in de methodesectie.

Tabel 1: Indicatoren voor ruimtelijke analyse

Indicator	Dimensie
Woonplaats / Werkplaats	Is het mogelijk woonplaats of werkplaats van respondent uit de data af te leiden?
Mate van lokaliteit	Beweegt een persoon zich vooral rondom een lokaal gebied, of komt het op vele plekken buiten de regio?
Tijdsduur	Hoelang is een respondent op de <i>stop</i> geweest?
Tijd van de dag	Op welke dagdelen beweegt een respondent zich?
Aantal provincies	In hoeveel verschillende provincies is een respondent geweest?
Frequentie	Hoe vaak komt de respondent op deze plek?

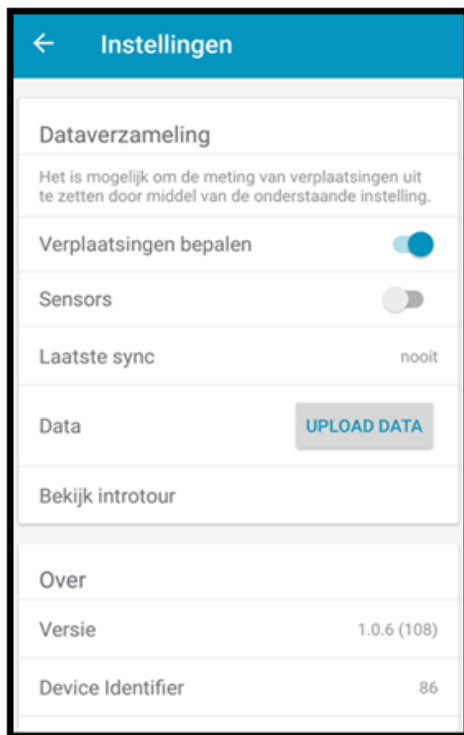
Methodologie

Binnen CBS is onderzoek met mobiliteitsdata nog vrij nieuw. Daarom heeft dit onderzoek ook een exploratief karakter. Omdat de Tabi-verplaatsing applicatie nog in de ontwikkeling is, is het in belang van dit onderzoek om te kijken naar de ruwe data en of daar nog meetfouten in zitten of uit voortkomen. Door de kwaliteit van de ruwe data te bestuderen kunnen er mogelijke fouten in de data en methodologie van de applicatie ontdekt worden en kan dit de ontwikkeling van de applicatie bevorderen. De ruwe data vormen immers de basis voor verdere visuele data. Deze visuele data zullen in dit onderzoek leidend zijn.

De data die gebruikt worden voor dit onderzoek zijn verzameld in november en december 2017. Bij deze meting deden 28 medewerkers (voornamelijk interviewers) vanuit het CBS – waarvan demografische kenmerken niet opgevraagd zijn – mee die uit verschillende omstreken van Nederland kwamen. Er is specifiek gekozen voor interviewers, omdat dit respondenten zijn die in verband met hun werk veel reizen. Dit maakt het een geschikte doelgroep om te testen of de Tabi-verplaatsing applicatie in verschillende regio's en binnen verschillende vervoersmiddelen de juiste metingen maakt.

De applicatie geeft respondenten de gelegenheid data door te sturen op ieder moment, waardoor een dynamische dataset ontstaat. Deze data wordt binnen CBS in een beschermde omgeving opgeslagen. Zoals te zien in **figuur 2** kunnen respondenten zelf bepalen op welk moment zij de data willen versturen. Zodra ze deze versturen is deze niet meer binnen de applicatie terug te vinden, dit om het geheugen van de telefoon niet teveel te belasten. Er is aan de participanten in dit onderzoek gevraagd of zij gedurende de testperiode de Tabi-verplaatsing applicatie zo vaak mogelijk actief kunnen laten draaien op hun smartphone, zodat de applicatie de kans krijgt zoveel mogelijk metingen te maken. Ook is er de participanten verzocht of ze de zogenoemde *stops* willen *labelen*. Door een *label* toe te voegen kunnen respondenten meer inhoud geven aan een locatie, waardoor de data meer toepasbaar wordt voor analyse. Respondenten zouden bijvoorbeeld in het *label* in kunnen vullen waarom zij ergens zijn.

Figuur 2: Respondent bepaalt wanneer data verstuurd wordt



Noot. Overzicht van het tabblad "instellingen" van de applicatie

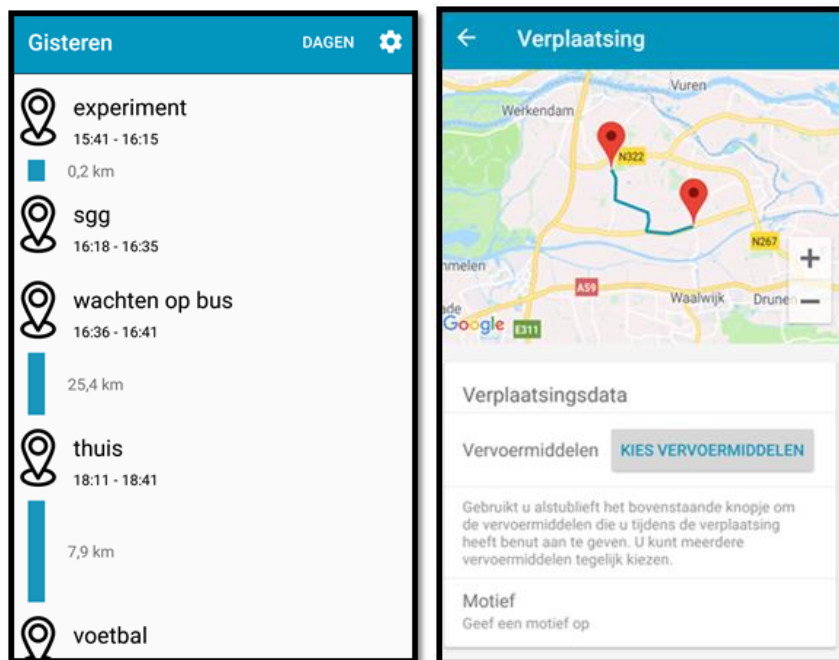
Alleen kwantitatieve metingen

In dit onderzoek is er bewust voor gekozen geen kwalitatief onderzoek te houden in de vorm van een interview met de respondenten. Een interview zou meer kennis geven over de context van een meetpunt of *stop* en geeft de mogelijkheid door te vragen over het doel van een respondent bij een bepaalde stop. Echter, het probleem bij een interview is, in het geval van mobiliteitsonderzoek, dat het voor de respondent lastig te herinneren is wat hij/zij bij een bepaalde meting aan het doen was; de reden waarom de Tabi-verplaatsing applicatie wordt ingezet. Zo herinneren respondenten zich niet altijd precies waar zij laatste week allemaal heen zijn gegaan en waarom zij dit deden (Chen et al., 2010). Wanneer zij dit deden wordt via de applicatie bijgehouden, echter het doel van hun mobiliteit is voor een respondent niet altijd meer juist op te halen. Er bestaat een kans dat hierdoor de data minder betrouwbaar wordt, doordat een respondent wellicht in een poging de situatie te herinneren verkeerde informatie verschaft.

Door het gevaar van afnemende betrouwbaarheid van de data houdt dit onderzoek de *stop* van respondenten als maatstaf. Hierbij is de optie om een *label* bij een *stop* in te vullen van belang, omdat dit inhoudelijke informatie en beschrijving voor de analyse bevat. Op deze manier kan de respondent het doel van een verplaatsing beschrijven en wordt het voor een

derde partij gemakkelijker om het te interpreteren. Verder kunnen respondenten hun *stops* en mobiliteitspatronen binnen de applicatie zelf terugkijken. Dit geeft hen de gelegenheid om eventueel op een ander tijdstip de *labels* in te vullen.

Figuur 3: Overzicht Tabi-verplaatsing applicatie



Noot. Deze screenshots zijn van een nieuwere versie van deze applicatie. Wel zijn deze in zekere mate vergelijkbaar met het scherm wat respondenten van deze *test-ronde* bij hun applicatie hebben gezien.

Nieuw wetenschappelijk inzicht op netwerken

Applicaties zoals de Tabi-verplaatsing zijn een alternatief voor de "klassieke" methode. Eerder is al gebleken dat tijdens het individueel registreren van verplaatsingen er een duidelijk verschil is binnen de kwaliteit van de data van automatische metingen en handmatige metingen (Hoogendoorn-Lanser, Schaap, Kalter & Coffeng, 2013). Doordat de Tabi-verplaatsing applicatie geautomatiseerd is, zullen problemen zoals het vergeten van een notitie vermeden worden, wat ten gunste komt voor de kwaliteit van de data. Deze geautomatiseerde methode zal waarschijnlijk meer valide resultaten opleveren. Daarbij vermindert geautomatiseerd onderzoek, buiten de meetfouten, ook de kosten van het onderzoek.

Dit onderzoek geeft een nieuw inzicht binnen het wetenschappelijk netwerk onderzoek: er kan hier namelijk analyse worden gedaan op basis van GPS-data. Het voordeel van deze methode is dat in plaats van het online netwerk, nu vooral uitspraak gedaan kan

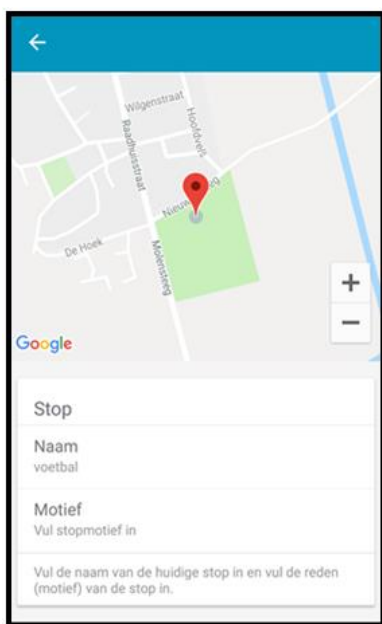
worden over het ruimtelijke netwerk. Het geeft een nieuwe methode weer die bij kan dragen aan netwerkonderzoek in het algemeen.

Ondanks dat de beschikbare data van de Tabi-verplaatsingen applicatie nog niet heel veel respondenten kennen, bieden deze al verschillende opties tot onderzoek. Data wat betreft mobiliteit, afkomstig uit de Tabi-verplaatsingen applicatie, zijn erg gedetailleerde data. Zo weergegeven deze vrij accuraat de positie van een respondent, is het mogelijk uitspraak over de hoeveelheid van de bezoeken van de respondent en bieden de data de respondent de kans om het doel van de mobiliteit aan te geven binnen de *label* van de *stop*. Bij dit alles moet wel vermeld worden dat deze gegevens alleen verzameld worden op het moment dat de applicatie geen meetfouten meer vertoont.

Hoe meet de applicatie?

De Tabi-verplaatsing applicatie maakt constant metingen als deze op de achtergrond op de telefoon actief is. Het interval tussen metingen is verschillend, afhankelijk van het besturingssysteem van je *smartphone*. Op het moment dat een respondent langer dan 5 minuten stil staat – dus geen fysieke beweging – verandert de status van de applicatie. Bevindt een persoon zich langer op eenzelfde plek dan meet de applicatie door het laatst gemeten punt nogmaals in te laden. Zodra een respondent buiten een zone van 200 meter buiten de laatst gemeten zone bevindt, start de applicatie weer met meten.

Figuur 4: Aan de hand van coördinaten bepaalt applicatie locatie



Noot. De optie tot *labeling* en het motief zijn beide optioneel in te vullen

Om de data te kunnen analyseren moet een vertaalslag gemaakt worden van de geproduceerde data van de Tabi-verplaatsing applicatie. De applicatie produceert aan de hand van verschillende sensoren ruwe data waarin tijd (tijd van meting), X en Y coördinaten (= LAT-LONG locatie), accuraatheid (welke sensor is er voor de meting gebruikt?) , user ID (identificeert respondent) en snelheid in km/uur bekend zijn (**zie tabel 2**). Hierbij is relevant om te noemen dat de accuraatheid van een meting afhankelijk is van de gebruikte sensor van het toestel. Over het algemeen zijn metingen met voldoende batterij op basis van GPS op ongeveer 10 tot 30 meter nauwkeurig. Deze data worden opgeslagen in een CSV-bestand binnen de servers van het CBS.

Tabel 2: Overzicht ruwe data

ID	Tijd van verzenden naar database	Latitude	Longitude	Snelheid (km/uur)	Accuraatheid	Afstand met vorige locatie	Gewenste sensor	Hoogte	Tijd van de meting
5	2016-16-26 04:26:47.344 542+01	5.219.23 5.474.42 8.950	5.130.577. 349.160.93 0	62	-1	0.214328261 6138458	-1	-1	2015-01-17 02:22:56.013 +01
6	2016-16-26 04:26:47.344 542+01	5.709.23 5.477.42 8.950	5.150.587. 449.060.63 0	55	-1	0.174328261 6138458	-1	0	2015-01-17 02:22:56.013 +01
7	2016-16-26 04:26:47.344 542+01	5.279.23 9.546.06 9.480	51.107.137 .723.092.6 08	15	-1	0.274128261 6138458	-1	9.815.573. 692.321.77 0	2015-01-17 02:22:56.013 +01

Noot. Fictieve data

Wat bij de verzamelde data van de applicatie ontbreekt is de visualisering van deze data. Met behulp van Rstudio zijn de verzamelde data per persoon geprojecteerd op kaartjes. Vanuit de ruwe data zijn uiteindelijk 28 kaartjes ontwikkeld. De ruwe kaartjes bestaan uit een kaart van Nederland met daarbij rode punten, wat staat voor "publieke gebouwen" en oranje punten, wat staat voor "infrastructurele elementen". Deze variabelen zijn in losstaande lagen toegevoegd. Op deze ruwe map zijn vervolgens de coördinaten uit de ruwe data geplot.

Vervolgens zijn de 28 geplote kaartjes gefilterd op basis van criteria die opgesteld zijn. Deze criteria worden later in de paragraaf toegelicht. Tijdens dit selectieproces is gelet op afwijkende, onwaarschijnlijke patronen en de hoeveelheid data op het kaartje. Hierbij kan gedacht worden aan patronen recht door water, data van slechts één dag of een datapatroon zonder *stops*. Na het filteren op grote afwijkende patronen en de hoeveelheid data is vastgesteld dat 16 van de 28 kaartjes relatief valide waren. Van deze 16 overgebleven kaartjes is nogmaals gekeken naar afwijkende patronen. Er is in dit selectieproces vooral gelet op het aantal *stops*, de hoeveelheid gemeten punten en de tijd die de respondent binnen dit onderzoek heeft meegedaan. Een kaartje is als ‘voldoende’ gedefinieerd wanneer het patroon aan de volgende voorwaarden voldoet. Ten eerste moest de respondent in zijn reispatroon een verschillend aantal plekken *stops* hebben gemaakt. Ten tweede moest het kaartje voldoende mobiliteitsdata bevatten, daarbij werden kaartjes met meer dan 5 dagen aan data goedgekeurd. Ten slotte moesten er geen (cruciale) meetfouten zichtbaar zijn op de kaart. Dit laatste criteria is niet geheel gewaarborgd, vanwege het feit dat sommige kaartjes, ondanks zichtbare meetfouten, wel waardevolle informatie bevatten. Uiteindelijk gebruikt dit onderzoek 5 kaartjes met informatieve, betrouwbare mobiliteitsdata.

Visuele informatie binnen de kaartjes

Vanuit de ruwe data zijn de kaartjes ontwikkeld. Rstudio biedt de mogelijkheid om de coördinaten te plotten op een landkaart. Dit programma lijkt de mobiliteitspatronen de plotten als een soort ‘laag’, waarin geen verdere context over het punt/coördinaat te zien is. Geavanceerde programma’s zoals Google Fusion Tables en ArcGIS bieden deze optie wel. Echter, vanwege praktische redenen is gekozen om binnen dit onderzoek met Rstudio te werken. Ook zou het ethisch onaanvaardbaar zijn om gevoelige data te verzenden naar Google, wat nodig is voor het programma Google Fusion Tables. Het gevolg hiervan is dat de visuele data - zogezegd de ontwikkelde kaartjes - minder gegevens weergeven dan er in werkelijkheid zijn. Zo is de variabele ‘tijd’ ook beschikbaar in de dataset, in dit onderzoek is deze variabele niet meegenomen omdat het via Rstudio te lastig was om te visualiseren. Ondanks dat de visuele kaartjes van dit onderzoek minder informatie bieden, lenen deze zich nog steeds voor analyse vanwege de zichtbare patronen van mobiliteit. Deze patronen zijn interessant voor het onderzoeken van netwerken. De ruwe gegevens zijn terug te vinden binnen het CSV-bestand met daarin de ruwe data van de Tabi-verplaatsing applicatie.

Tekortkomingen methodologie en data van Tabi-verplaatsingen

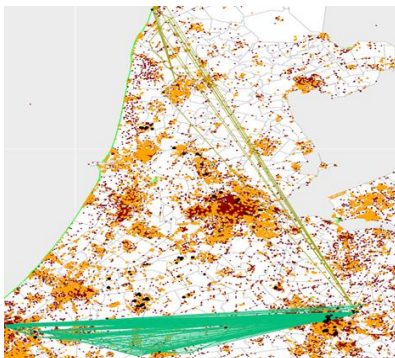
16 kaartjes voldeden aan de milde opgezette betrouwbaarheids- en validiteitscriteria en slechts 5 kaartjes bieden voldoende data met heldere mobiliteitspatronen. Bij de overige 23 kaartjes werden nog verschillende fouten in de metingen teruggevonden op de gevisualiseerde kaartjes. De oorzaak van de foutieve data is niet altijd eenduidig, er kunnen meerdere oorzaken zijn voor de foutieve data. De meetfouten kunnen onder andere door het handelen van de respondent veroorzaakt worden (bijvoorbeeld: respondent vergeet applicatie aan te zetten of batterij van *smartphone* is leeg), maar ook kunnen de meetfouten een software/*smartphone* gerelateerde oorzaak hebben. Zo is bekend dat tijdens de eerste *testronde* de applicatie op sommige modellen telefoons vaak zichzelf op inactief zetten en is de gebruikte sensor voor een meting (Wi-fi positionering/GPS) deels afhankelijk van het besturingssysteem van de telefoon (Android of IOS).

Foutieve patronen: meetfouten

Zoals eerder aangegeven stuurt de Tabi-verplaatsing applicatie de ruwe data door naar een bestand op een webserver binnen het CBS. Met dit bestand zijn de GPS-coördinaten geplot op de kaartjes via Rstudio. Hieronder staat een voorbeeld van een kaartje waarbij opvallende en implausibele patronen afgebeeld zijn. Er zijn meerdere afwijkende patronen gevonden die tot invalide kaartjes hebben geleid. Verschillende van deze zijn terug te vinden binnen de data van **figuur 5A**. Ten eerste de rechte lijnen tussen twee meetpunten. Deze data geven aan dat een respondent in een rechte lijn heeft gereisd. Dit is onwaarschijnlijk, omdat **figuur 6** laat zien dat dit niet overeenkomt met de Nederlandse infrastructuur.

Fout 1: "Waaier-patroon"

Figuur 5A: Respondent lijkt reis onnodig veel te maken



Noot. Ruwe GPS-data van de applicatie geplot via Rstudio

Verder wordt duidelijk uit deze figuur dat op één dag heel veel is gereisd tussen drie punten. Ondanks dat het vele op en neer reizen niet in strijd is met de bekende natuurwetten, lijkt het erg onwaarschijnlijk dat een respondent zo vaak op en neer reist tussen drie dezelfde punten op eenzelfde dag. Waarschijnlijk is deze respondent slechts één of twee keer naar Den Haag geweest via Gouda. Daarbij zou het ontwikkelde mobiliteitspatroon in **figuur 5B** qua tijd ook niet haalbaar zijn. Dit is een foutief patroon wat op meerdere kaartjes binnen de analyse is teruggevonden. Dit patroon maakt het moeilijk voor onderzoekers om te achterhalen wat er nu werkelijk afgereisd is binnen deze dag.

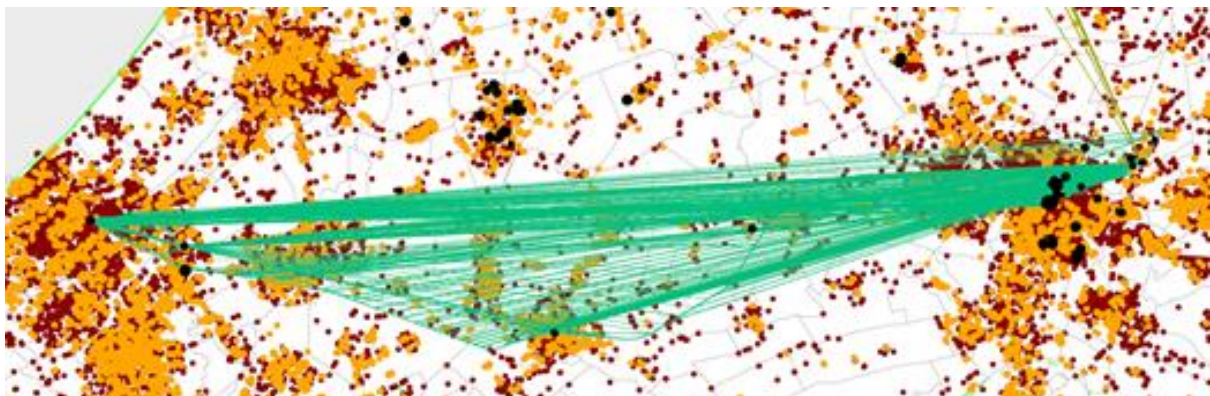
Figuur 6: Overzicht Nederlandse infrastructuur



Bron. Tuxx. (z.d.).

Noot. Alleen snelwegen

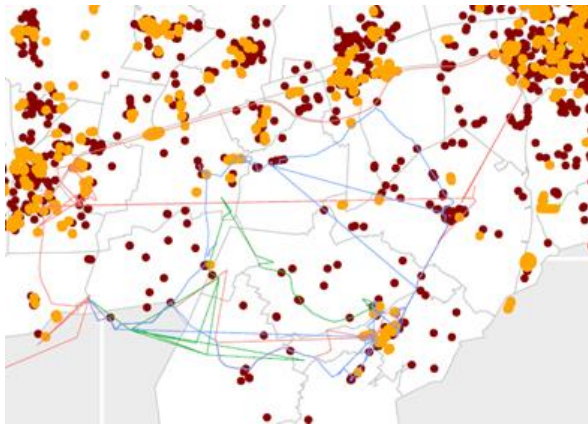
Figuur 5B: Foutief reispatroon van dichtbij



Naast het 'waaier' patroon van hierboven zijn ook verschillende respondenten welke weinig tot geen data hebben verzameld. Zo suggereert **figuur 7** dat er bij sommige respondenten erg weinig en onduidelijke metingen zijn gemaakt. Te zien is dat de respondent zich heeft verplaatst, zonder hierbij op een punt te stoppen. Aan de hand van dit kaartje kan gesteld worden dat een respondent op verschillende dagen reist, maar dat de desbetreffende respondent daarbij geen *stops* maakt. Dit lijkt een erg onwaarschijnlijk patroon, waardoor dit kaartje zich niet leent als valide kaartje voor dit onderzoek. Een gebrek aan *stops* heeft ook tot gevolg dat niet duidelijk wordt waar een respondent zich vooral bevindt.

Fout 2: Ontbrekende data

Figuur 7: Een gebrek aan data van de respondent



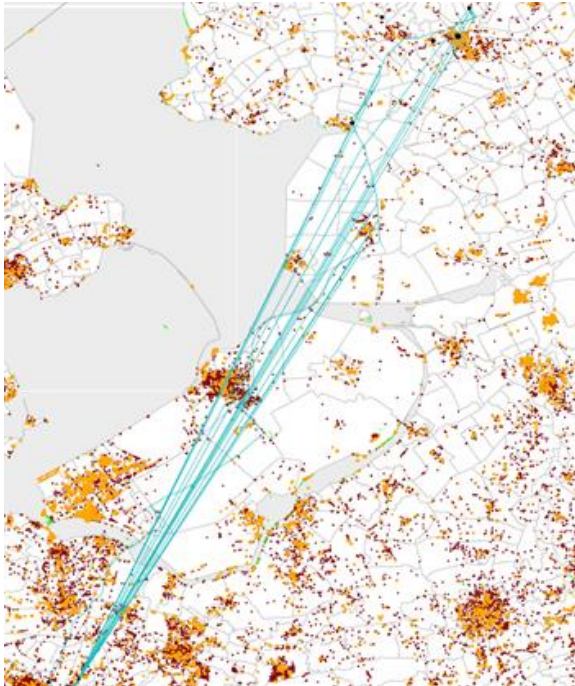
Noot. Reispatroon in Zuid-Nederland zonder enige 'stop'

De kaartjes laten meerdere keren rechte lijnen tussen verschillende metingen zien. Buiten dat dit fenomeen op zichzelf al invalide is, zorgt het ook voor mobiliteitspatronen recht door het water, wat onwaarschijnlijk is. Binnen **figuur 8** wordt het patroon op een manier afgebeeld dat een respondent recht door het IJsselmeer reist. In realiteit is het aannemelijker dat de respondent gewoon over het land naar zijn/haar bestemming is gegaan. Doordat er in dit kaartje alleen rechte lijnen te zien zijn is het niet mogelijk vast te stellen van wat voor wegen de respondent gebruik heeft gemaakt.

Ook is er binnen dit kaartje slechts data te zien van één dag. Misschien was deze dag wel een weekenddag, een dag waarop de respondent misschien wel hele andere bestemmingen heeft dan op een doordeweekse dag. Om een goed beeld te krijgen is binnen dit onderzoek een vereiste dat een kaartje data laat zien van tenminste 5 dagen.

Fout 3: Meetpatronen over water

Figuur 8: Lijnen recht door water



Noot. Respondent reist recht daar het water in plaats van over het land

Foutieve patronen: oorzaken

De Tabi-verplaatsing applicatie is afhankelijk van sensoren die de metingen doen. Het is belangrijk dat deze metingen tijdig en consequent gedaan worden, de applicatie verbindt namelijk twee opeenvolgende meetpunten met elkaar in een rechte lijn. De kaarsrechte lijnen, welke op verschillende kaartjes terug te vinden zijn, lijken zo erg op een gevolg van een wegvallend signaal van de sensoren van de applicatie. Op het moment dat de applicatie het GPS-signaal kwijt is, wacht deze tot er weer een nieuwe meting is en deze punten worden in een rechte lijn met elkaar verbonden. De wegvallende data kan zo ook de reispatronen recht door het water verklaren. Blijkbaar is de applicatie in de tussentijd het GPS-contact met de respondent kwijt geweest. Met de data waar dit in voor gekomen is heeft de respondent de getoonde route niet precies afgelegd. Buiten de directe gevolgen (rechte lijnen, reispatronen door water) is een verder gevolg dat de route welke in realiteit door de respondent is afgelegd onbekend blijft.

Het ‘waaier’ patroon zoals in de **figuren 5A** en **5B** blijft tot op heden een lastig te verklaren patroon. Mogelijke oorzaken kunnen zijn het ingelogd zijn van een respondent op verschillende toestellen, wat een signaal constant op en neer laat springen. Ook kan het zijn dat tijdens de bewerking van de data de GPS-coördinaten door elkaar zijn geraakt. Vanwege

het feit dat elk toestel een losse *device code* heeft kan uitgesloten worden dat verschillende toestellen - respondent is bijvoorbeeld ingelogd op zijn telefoon en op zijn Ipad - in eenzelfde figuur worden afgebeeld. Ook is het CSV-bestand van de data gecheckt op de volgorde van de data. Hieruit bleek dat data op volgorde gesorteerd stond binnen dit bestand, wat de tweede mogelijke verklaring van het "waaier" patroon weerlegt.

Een gebrek aan data kan verklaard worden aan de hand van het gedrag van de respondent. Aangezien dit de eerste test was met de applicatie was het misschien voor alle respondenten niet duidelijk dat zij de applicatie altijd actief moet houden op een tweede scherm. In sommige gevallen dient de applicatie ook weer actief te worden gezet door op het icoon van de applicatie te drukken. Respondenten die niet goed weten wat van hen verwacht wordt hebben een grote kans op het maken van dit soort fouten. Ook is voor respondenten niet duidelijk gezegd hoe lang zij de applicatie moeten laten draaien, waardoor sommige mensen de applicatie wellicht maar 2 dagen aan hebben laten staan. Uit feedback op de applicatie is wel gebleken dat gedurende deze test het actief hebben aanstaan van de telefoon zwaar meeweegt in de duur van de batterij en het geheugen van een telefoon. Dit kunnen beide redenen zijn voor een respondent om de GPS-sensor uit te schakelen, waardoor de applicatie het signaal óf kwijt is óf minder accuraat kan weergeven.

Alle bovengenoemde fouten kunnen worden gekoppeld aan methodologische of ICT-gerelateerde fouten. Beide dimensies worden gebruikt om de fouten binnen de data te verklaren. In onderstaand overzicht zijn alle oorzaken voor afwijkingen beschreven en hierbij is aangegeven welke dimensie hierbij van toepassing is. Een overzicht van de meetfouten die van invloed zijn geweest op de ruwe data vindt u in de onderstaande **tabel 3**.

Tabel 3: Overzicht oorzaken foutieve patronen

<i>Meetfout door handelen respondent</i>	<i>Meetfout binnen software Tabi-verplaatsing applicatie</i>
Geen GPS-signaal van respondent waardoor er geen metingen worden gemaakt	Wel signaal maar onnauwkeurig (enkel GSM)
Respondent heeft het toestel waarop de applicatie geïnstalleerd is niet meegenomen gedurende de verplaatsing	Applicatie is vastgelopen en meet tijdelijk niet
De batterij/accu van het toestel is leeg (geraakt) door de dag heen.	Applicatie detecteert te laat beweging en heeft daardoor een te lange rustperiode
Respondent is ingelogd op verschillende toestellen tegelijkertijd	<i>Snapping</i> : Dit probleem kunnen we oplossen door te kijken naar de snelheden die zijn afgelegd. Ook moeten we kijken naar de <i>accuracy</i> van de meetpunten en of die veranderen (hoe zit het met de straal van het meetpunt?) <i>GPS offset</i> : Dit komt voor in gebieden met hoge gebouwen en veel bos. Om dit op te lossen zouden we in een gebied moeten rondlopen met een duidelijk gps signaal en achteraf kijken of de aangegeven route in de app overeenkomt met de daadwerkelijke gelopen route.

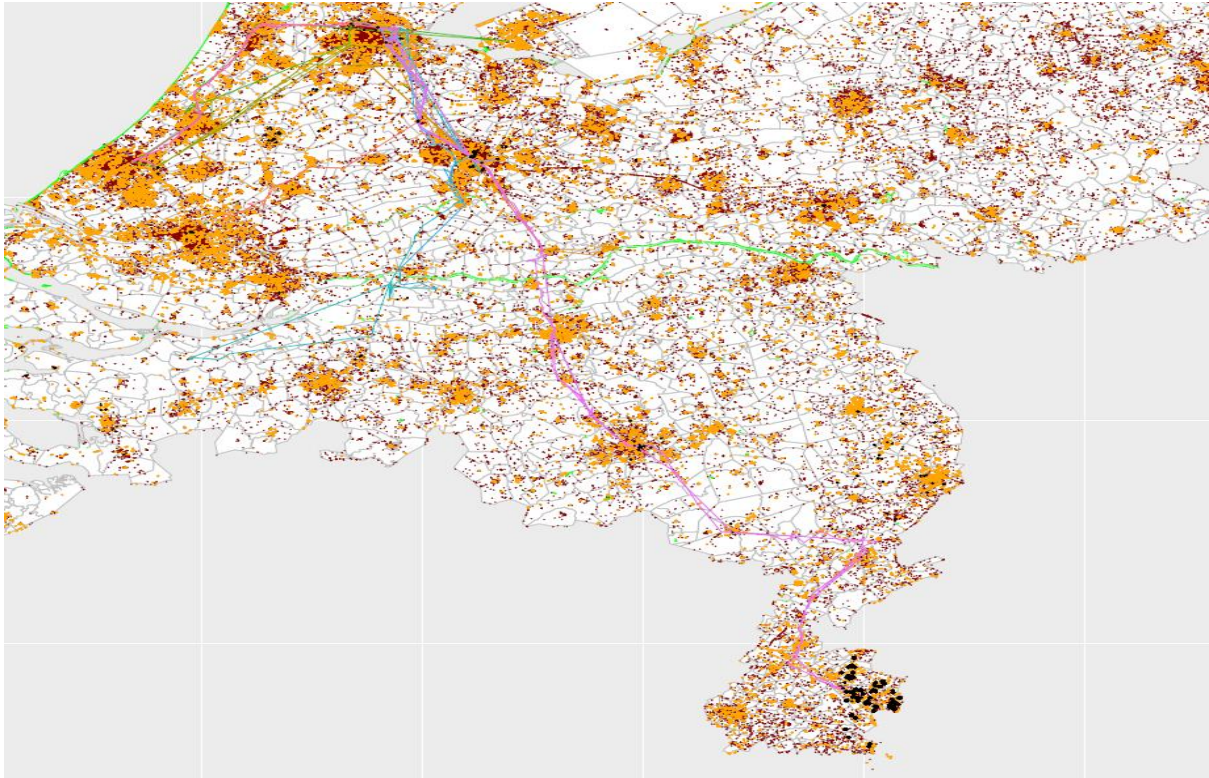
Resultaten

Analyse van de mobiliteitsdata

Respondent 6 woont in Amsterdam, aangezien deze stad het centrale vertrekpunt is waar de meeste verplaatsingen vandaan gaan(zie: **figuur 9**). Hij/zij reist vaak met de trein naar het CBS in Den Haag, dit zal waarschijnlijk de werkomgeving van de respondent zijn. Er is vaak via Utrecht Centraal naar Eindhoven en vervolgens naar Heerlen gereisd en er zijn rondom Amsterdam verschillende uitstappen gemaakt naar omliggende gebieden. Ook de verplaatsingen naar Heerlen lijken werk gerelateerd, dit valt af te leiden uit het feit dat de doelgroep van onze steekproef CBS-medewerkers zijn en ook in Heerlen een afdeling van het CBS zit. Zowel in Den Haag, Amsterdam, Utrecht, Eindhoven als in Heerlen heeft deze respondent een *stop* gemaakt. In dit geval komt het door het wachten op de trein en bezig zijn op een vaste werkplek. Uit **figuur 9** blijkt dat respondent 6 in vijf verschillende provincies geweest, dit toont aan dat de respondent zich veel verplaatst en een verspreid sociaal netwerk lijkt te hebben. In totaal heeft respondent 6 twintig *stops* gemaakt. Veel van de *stops* zijn gemaakt Amsterdam en in Heerlen wat het idee van de woon- en werkplek ondersteund. De respondent beweegt veel zowel doordeweeks als in het weekend.

Het is aannemelijk dat de respondent een sociaal netwerk heeft dat verspreid zit over vier grote steden. Ten eerste Amsterdam, waar respondent 6 waarschijnlijk woont en veel persoonlijke contacten kan hebben. De overige drie locaties zijn locaties waar de respondent waarschijnlijk voor werk naar toe gaat, de contacten op deze punten zullen vooral collega's zijn. Te zien is ook dat in Heerlen de meeste *stops* zitten. Dit zou kunnen betekenen dat de relaties die hij daar heeft sterker zijn dan in Den Haag, omdat respondent 6 daar meer tijd heeft besteed en sterkere banden heeft kunnen ontwikkelen.

Figuur 9. GPS-data van verplaatsingen van respondent 6 van 11 november tot 12 december

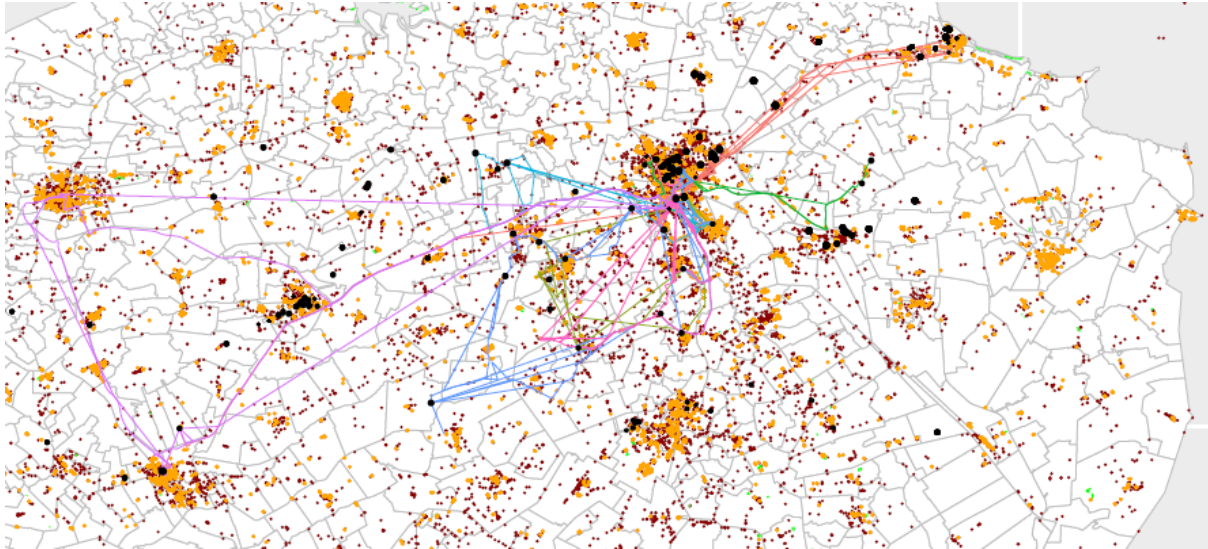


Respondent 10 (**figuur 10**) heeft een duidelijke werk- en leefomgeving die zich in Groningen bevindt. Op het kaartje is duidelijk terug te zien dat de respondent de meeste tijd doorbrengt in de provincie Groningen. Zijn/haar werkplaats bevindt zich waarschijnlijk in Groningen, omdat daar de meeste *stops* zijn gemaakt. Respondent 10 heeft ook verschillende uitstapjes gemaakt binnen de provincie Groningen, dit kunnen triviale activiteiten zijn zoals naar de supermarkt of naar vrienden toe reizen. Respondent 10 heeft in totaal 32 *stops* gemaakt. Grotendeels van deze *stops* zijn gemaakt op de woonplek van respondent 10. Dit suggereert dat respondent 10 vaak thuis zit en je zou hierdoor een sociaal netwerk verwachten die een sterk lokaal karakter heeft. Ook deze respondent beweegt zich vaak in doordeweeks en in het weekend.

Ook wordt duidelijk dat respondent 10 enkele keren naar Leeuwarden en andere steden in Friesland is geweest. Dit kunnen werk gerelateerde activiteiten zijn of het bezoeken van vrienden en familie. Het is duidelijk te zien aan de hand van onderstaand **figuur 10**, dat het sociale netwerk van de respondent erg gecentreerd zit in Groningen en gedeeltelijk in Friesland. Dit zijn de enige twee provincies waar respondent 10 is geweest in de afgelopen

weken. Hiermee kan gezegd worden dat het netwerk van respondent 10 voornamelijk lokale vrienden en familie zijn.

Figuur 10. GPS-data van verplaatsingen van respondent 10 van 29 november tot 21 december



Wat meteen opvalt in **figuur 11** is dat respondent 11 op veel verschillende locaties is geweest. Het is niet duidelijk vast te stellen wat precies de woon- en werkplek is van respondent 11. Dit komt doordat er in verschillende grote steden veel verplaatsingen plaatsvinden. Normaal kan er uit de kaartjes één duidelijk vertrekpunt afgeleid worden, echter in dit geval niet.

Verder lijkt evident dat veel verplaatsingen vanaf Utrecht en Amsterdam naar grote steden zoals Den Haag, Groningen en Assen plaatsvinden. Respondent 11 maakt in totaal 33 stops gedurende de periode. Opmerkelijk is dat de meeste stops gemaakt zijn in Utrecht en in Utrecht ook meer frequent dezelfde stops zijn gemaakt. Dit zou bewijs kunnen zijn voor het feit dat respondent 11 in Utrecht zou wonen, omdat hij/zij daar de meeste tijd doorbrengt. Respondent 11 reist veel doordeweeks en is in het weekend minder actief.

Wat opvalt aan **figuur 11**, is dat er lijnen over het IJsselmeer gaan. Deze kloppen niet en zouden allemaal via de E22 naar Friesland en Groningen moeten gaan, omdat dat de enige mogelijke weg is. Ondanks deze afwijking biedt dit kaartje wel veel informatie over de reispatronen van respondent 11. Te zien is dat respondent 11 een turbulent leven heeft waarin er veel gereisd wordt. Er worden regelmatig lange afstanden afgelegd die waarschijnlijk werkgerelateerd zijn. Respondent 11 is in totaal in vijf verschillende provincies geweest.

Respondent 11 heeft waarschijnlijk een divers sociaal netwerk omdat er veel plekken zijn bezocht op regelmatige basis. Ook zijn er enkele uitstapjes gemaakt naar Haarlem, dit geeft aan dat respondent 11 ook op een lokaal niveau een sociaal netwerk heeft (bijvoorbeeld: bezoeken van vrienden en familie).

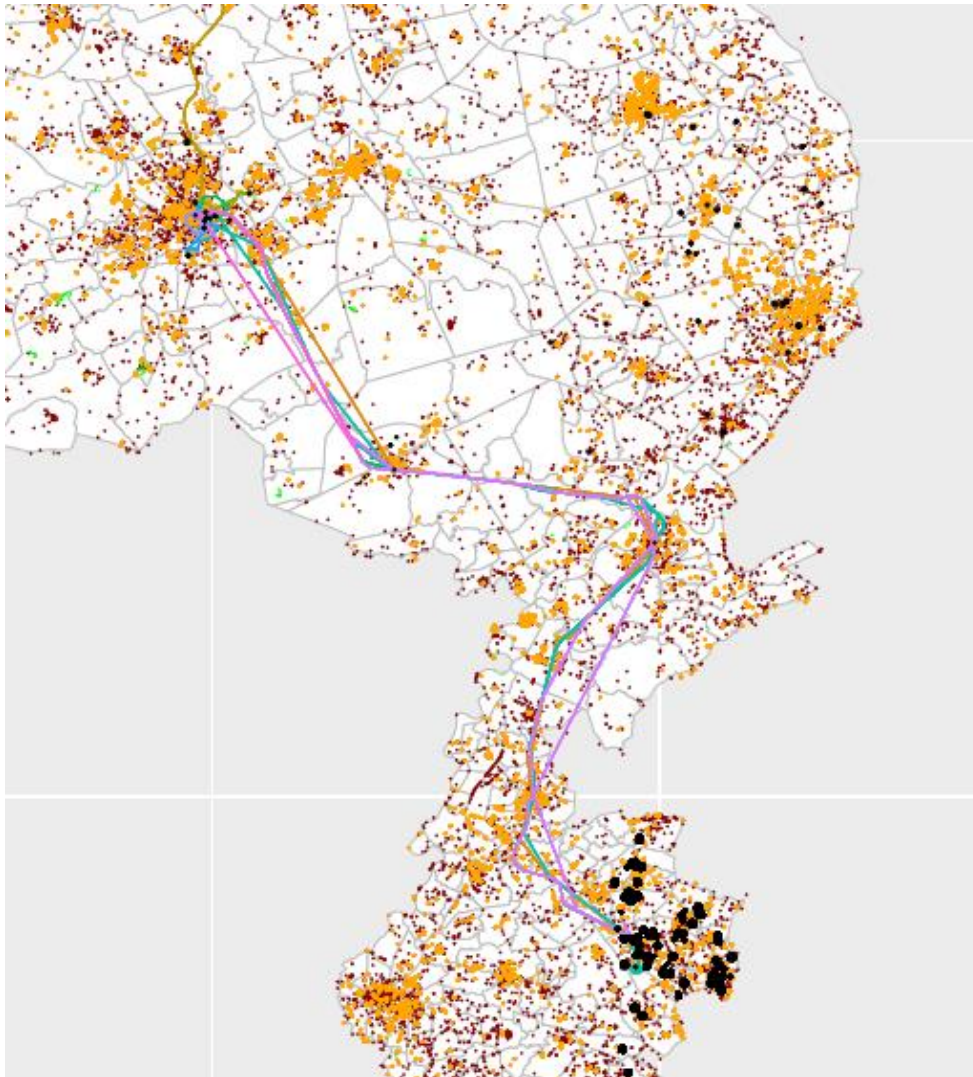
Figuur 11. GPS-data van verplaatsingen van respondent 11 van 29 november tot 21 december



In **figuur 12** vindt u een illustratie van een respondent met een duidelijk mobiliteitspatroon. Respondent 17 woont in Eindhoven en heeft over verschillende dagen dezelfde route naar Heerlen heeft genomen. Dit zijn dan ook de enige twee locaties waar de respondent zich de laatste periode heeft bevonden en waar ook de meeste *stops* te zien zijn. Respondent 17 maakt in totaal 26 stops. Meeste van deze stops bevinden zich rondom Eindhoven en in Heerlen. In Heerlen zie je duidelijk dat daar vaker stops op dezelfde plek zijn gemaakt. Dit ondersteunt het idee dat respondent 17 in Eindhoven woont en in Heerlen werkt. Respondent 17 reist veel doordeweeks, in het weekend verplaatst hij/zij zich amper en is voornamelijk thuis in Venlo.

Het sociale netwerk van respondent 17 is beperkt tot zijn/haar werkplek en daarom ook minder divers omdat zijn/haar relaties voornamelijk uit CBS medewerkers bestaat. De respondent is in twee provincies geweest. Wat aangeeft dat deze persoon redelijk lokaal leeft. Dit is een duidelijk voorbeeld van een respondent die reist voor noodzakelijke behoeften, in dit geval zijn/haar werk.

Figuur 12. GPS-data van verplaatsingen van respondent 17 van 6 december tot 3 maart

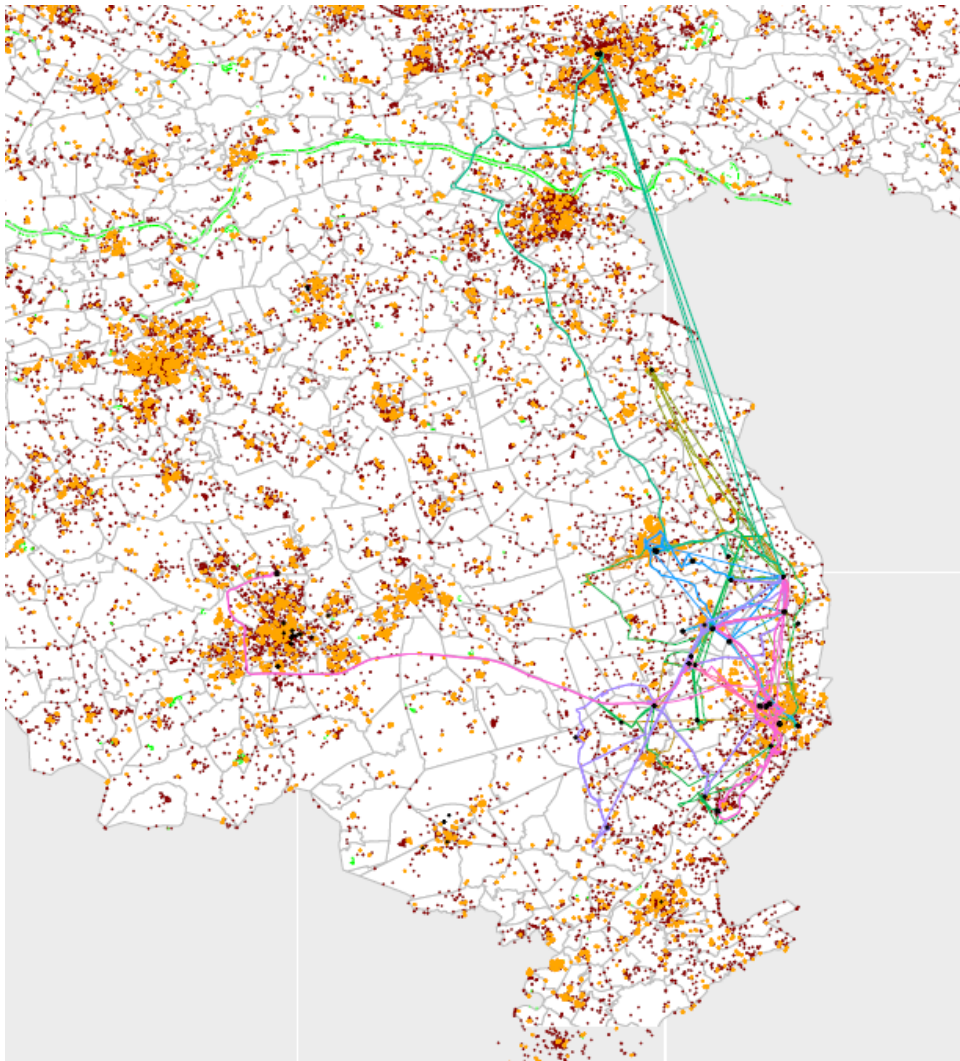


Uit de patronen in **figuur 13** is de woonplek van respondent 21 af te leiden. Deze bevindt zich waarschijnlijk in Venlo. Dit is het centrale punt waar vanuit de meeste verplaatsingen plaatsvinden. Respondent 21 heeft of een baan in Venlo of werkt vooral thuis, omdat er verder niet veel reispatronen naar één bepaalde locatie gaan. De tweede optie is meer waarschijnlijk, omdat de respondent in dit geval een CBS-medewerker is en het CBS in Den Haag en Heerlen zitten. In totaal maakt respondent 21 achtendertig stops waarvan veel stops

rondom Venlo liggen. Dit ondersteunt het idee dat respondent 21 in Venlo woont. Respondent 21 reist vooral doordeweeks en in het weekend weinig.

Verder is respondent 21 een enkele keer naar Arnhem geweest. Buiten dat wordt ook duidelijk dat de respondent veel rondom zijn woonplek reist. Hij legt geen lange afstanden af en is daarom ook maar in twee provincies geweest. Hieruit kan opgemaakt worden dat respondent 21 een lokaal sociaal netwerk heeft. De mobiliteitsdata suggereren dat hier, omdat veel *stops* te zien zijn in omliggende gebieden. Dit geeft aan dat respondent 21 daar een langere tijd heeft stilgestaan (bezoeken van vrienden of activiteiten).

Figuur 13. GPS-data van verplaatsingen van respondent 21 van 14 december tot 10 januari.



Onderstaande **tabel 4** vat alle geobserveerde indicatoren die in de resultatensectie besproken zijn samen.

Tabel 4: overzicht resultaten

Respondent	Woonplaats	Werkplaats	Lokaal/Landelijk	Frequentie	Aantal <i>stops</i>	Reist veel doordeweeks, weekend of allebei
6	Amsterdam	Den Haag/Heerlen	Landelijk	Vooraf veel <i>stops</i> op plekken voor werk	20	Allebei
10	Groningen	Groningen	Lokaal	-	32	Allebei
11	Utrecht/Amsterdam	-	Landelijk	-	33	Doordeweeks
17	Eindhoven	Heerlen	Lokaal	Vooraf veel <i>stops</i> op plekken voor werk	26	Doordeweeks
21	Venlo	-	Lokaal	Vooraf in de buurt van woning	38	Doordeweeks

Conclusie & discussie

Dit onderzoek heeft een nieuw inzicht gegeven binnen de sociale netwerk analyse waarvan gebruik is gemaakt van GPS-data die fysieke verplaatsingen aangeeft. Het bestuderen sociale netwerken is interessant en relevant omdat het helpt met het verklaren van sociale fenomenen (bijvoorbeeld: vinden van een baan). De hoofdvraag die centraal staat in ons onderzoek is wat voor een bijdrage de Tabi-verplaatsingen applicatie levert aan sociaal netwerkonderzoek en wat er afgeleid kan worden uit deze gegevens.

Wat speciaal is aan de tabi-verplaatsingen applicatie is dat het de fysieke aspect belicht van netwerkonderzoek en er metingen gedaan worden naar bewegingspatronen. Dit is een andere vorm van data dan dat normaal gebruikt wordt. Normaal wordt er gebruik gemaakt van online data van social media of vragenlijsten om zo het sociaal netwerk van een individu te meten. GPS-data zou het offline netwerk van een individu kunnen meten, wat een tekortkoming is bij online netwerkonderzoek.

Zo zijn er binnen ons onderzoek enkele indicatoren opgesteld die kunnen worden afgeleid uit de GPS-data. Alle opgestelde indicatoren konden in bijna alle gevallen verklaard worden aan de hand van de GPS-data. Zowel de woonplaats als de lokaliteit van het netwerk konden bij alle respondenten van de kaartjes afgeleid worden.

In onze resultatensectie is naar voren gekomen dat men vaak vanuit hun woonplaats reist naar hun werk. Uit de plaatjes kan er geconcludeerd worden dat herhalende activiteiten zoals werk zorgen voor een vaste routine die duidelijk terug is te zien in het reispatroon van de respondent. Daarnaast is er geprobeerd uitspraken te doen over het sociale netwerk van de respondent. Mensen die zich in meer dan 3 provincies hadden verplaatst werden gecategoriseerd als een individu met een divers netwerk. Ook is er gekeken naar het aantal *stops* van een individu. Alle respondenten hadden 20 of meer stops gedurende de periode van metingen. Deze stops waren vaak gemaakt op de woon of werkplek van de respondent. Tenslotte is er gekeken naar in welke deel van de week er het meest gereisd wordt. De resultaten waren redelijk naar verwachting. Wij dachten meer reisactiviteit aan te treffen doordeweeks dan in de weekenden omdat onze respondenten allemaal een baan hebben. Drie van de vijf respondent verplaatste zich het meest doordeweeks, de overige twee respondenten verplaatste zich zowel vaak doordeweeks als in het weekend.

Op basis van onze theorie verwachten wij dat hoe verder een respondent reist, hoe minder sociale relaties hij zal hebben (Musolesi, 2007). Resultaten van dit onderzoek weerleggen dit idee. Op basis van de kaartjes blijkt dat respondenten die lange afstanden

afleggen veel *stops* maken en dus ook op lange afstanden een sociaal netwerk hebben. Dit kan echter wel verklaard worden door het feit dat CBS medewerkers veel reizen en in zowel Den Haag als in Heerlen werkzaam zijn. Verder wordt er gesteld dat vaste routines (werk, school) en sociale relaties een vast bewegingspatroon creëren (Belot & Ermisch, 2009). Onze resultaten tonen aan dat werk een vast reispatroon creëert. Er zijn op verschillende dagen herhalende reispatronen te zien naar dezelfde werkplek. In diezelfde zin zou de werkplek van een respondent ook zijn sociale netwerk kunnen zijn en daardoor ook zorgen voor een vast reispatroon. Mobiliteitspatronen van mensen zijn tijd en locatie gebonden. Dit betekent dat rond bepaalde tijden vaak dezelfde locaties worden gezocht en dat men frequent terugkeert naar deze locaties (Gonzalez et al., 2008). Deze theorie hebben we binnen ons onderzoek niet kunnen bevestigen omdat we de variabele "tijd" niet hebben meegenomen omdat die variabel niet te visualiseren was. Zou de variabel tijd wel meegenomen zijn, dan zou er hoogstwaarschijnlijk wel ondersteuning voor de theorie gevonden zijn. Dit komt doordat we vaste routines hebben binnen ons leven die ervoor zorgen dat we ons elke dag op dezelfde tijden verplaatsen (Cho et al., 2011).

Een beperking binnen dit onderzoek is dat het lastig is geweest objectief de kaartjes te bestuderen. Zo zijn, doordat duidelijk was dat het om CBS-medewerkers ging, vooraf al karakteristieken van de respondent bekend geweest. Tijdens het analyseren was zo duidelijk dat de werkplek van de respondent logische wijze in één van de filialen van het CBS zou moeten zijn (Den Haag of Heerlen).

Ook bij de data en de visualisering ervan waren er enkele tekortkomingen. Idealiter hadden de respondenten aan hun *stops* een duidelijke label gegeven, waardoor het interpreteren en het begrijpen van de context makkelijker zou worden. Echter, omdat bij het verzamelen van de data niet de instructie is gegeven om *stops* te *labelen* hebben respondenten veel *stops* niet beschreven. Dit maakte de analyse moeilijker. Daarnaast gaven onze kaartjes niet de volledige infrastructuur en bebouwing van Nederland. De aanwezigheid hiervan zou aangeven in wat voor soort gebouwen de respondent zijn tijd doorbrengt en een positieve bijdrage leveren aan het sociale netwerk onderzoek. Dit was niet mogelijk omdat dit toch te grote bestanden zou leiden in Rstudio, wat het onderzoek praktische beperkingen zou opleggen.

Om vervolgonderzoek te doen naar dit onderwerp moet er zoveel mogelijk gedaan worden aan de tekortkomingen die hierboven beschreven zijn. Wat vooral belangrijk is is dat we moeten kunnen onderscheiden wanneer iemand reist voor zijn werk of voor het bezoeken van zijn sociale netwerk. Dit kan worden gedaan door het aangeven van verschillende soorten

gebouwen en huizen op zo'n kaart. Daarnaast kan de respondent ook helpen in dit onderzoek, door aan de *stops* een duidelijke label te geven, dit maakt het interpreteren van de data makkelijker omdat het duidelijk is wat de respondent tijdens die *stop* gedaan heeft. Dit kan gedaan worden door de respondenten voor het afnemen van het onderzoek instructies te geven om de *stops* te labelen. Ook is het handig om in de applicatie voorafgaand een vragenlijst op te stellen waarin gevraagd kan worden of hij een rijbewijs heeft, of diegene een baan heeft en wat voor voertuigen hij bezit. Alle relevante data die je vooraf verzameld van een respondent maakt ook het interpreteren en analyseren een stuk makkelijker. Onze laatste aanbeveling is om een random steekproef te nemen. In het huidige onderzoek is er vooral gekozen voor CBS-medewerkers om praktische redenen. Bewegingspatronen kunnen hierdoor overlappen omdat ze bij dezelfde organisatie werken. Dit kan zorgen voor een vertekend beeld. Willekeurige selectie van respondenten zorgt voor een hogere betrouwbaarheid van de data die ook representatief is voor de populatie. Kortom is het gebruik van GPS-data een waardevolle manier om sociale netwerken te bestuderen en te begrijpen als bovenstaande gebreken van de huidige Tabi-verplaatsingen applicatie verholpen worden.

Referenties

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall.

Belot, M., & Ermisch, J. (2009). Friendship ties and geographical mobility: evidence from Great Britain. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(2), 427-442.

Bianchi, S. M., Milkie, M. A., Sayer, L. C., & Robinson, J. P. (2000). Is anyone doing the housework? Trends in the gender division of household labor. *Social forces*, 79(1), 191-228.

Bureau of Labor Statistics. (2016). Average hours per day spent in selected activities by employment status and sex. Geraadpleegd op 28 maart 2018, van <https://www.bls.gov/charts/american-time-use/activity-by-emp.htm>

Chen, C., Gong, H., Lawson, C., Bialostozky, E. (2010) Evaluating the feasibility of a passive travel survey collection in a complex urban environment: lessons learned from the New York City case study. *Transportation Research*, A(44), 830-840.

Cho, E., Myers, S. A., & Leskovec, J. (2011, August). Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1082-1090). ACM.

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. *Journal of computer-mediated communication*, 12(4), 1143-1168.

Feld, S. L. (1981). The focused organization of social ties. *American journal of sociology*, 86(5), 1015-1035.

Gao, H., Tang, J., & Liu, H. (2012, June). Exploring social-historical ties on location-based social networks. In *Icwsn*.

Hamermesh, D. S., Frazis, H., & Stewart, J. (2005). Data watch: The American time use survey. *Journal of Economic Perspectives*, 19(1), 221-232.

Het Kennisinstituut voor Mobiliteitsbeleid. (2017). Voor alle vormen van mobiliteit. Geraadpleegd op 16 april 2018, van <https://www.kimnet.nl/over-het-kim>

Hoogendoorn-Lanser, S., voor Mobiliteitsbeleid, K., Schaap, N., Kalter, M. J. O., & Coffeng, G. (2013). Het Mobiliteitspanel Nederland: een oude bekende in een nieuw jasje. In *Colloquium Vervoersplanologisch Speurwerk*.

J. Goldenberg and M. Levy. Distance is not dead: Social interaction and geographical distance in the internet era. Arxiv, abs/0906.3202, 2009.

Jack, S. L., & Anderson, A. R. (2002). The effects of embeddedness on the entrepreneurial process. *Journal of business Venturing*, 17(5), 467-487.

Liben-Nowell, D., Novak, J., Kumar, R., Raghavan, P., & Tomkins, A. (2005). Geographic routing in social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(33), 11623-11628.

M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, 2008.

Musolesi, M. & Mascolo, C. (2007). Designing mobility models based on social network theory. *ACM SIGMOBILE Mobile Computing and Communications Review*, 11(3), 59-70.

N. Eagle and A. Pentland. Eigenbehaviors: Identifying structure in routine. *Behavioral Ecology and Soc.*, 2009.

Eagle, N., Pentland, A. S., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the national academy of sciences*, 106(36), 15274-15278.

Passy, F., & Giugni, M. (2001, March). Social networks and individual perceptions: Explaining differential participation in social movements. In *Sociological forum* (Vol. 16, No. 1, pp. 123-153). Kluwer Academic Publishers-Plenum Publishers.

Tajfel, H., & Turner, J. C. (2004). *The social identity theory of intergroup behavior*.

Tijhuis, M. A. R., Flap, H. D., Foets, M., & Groenewegen, P. P. (1992). Netwerken in Nederland. Een onderzoek naar persoonlijke netwerken van Nederlanders. *Mens en maatschappij*, 67(1), 5-22.

Toole, J. L., Herrera-Yañe, C., Schneider, C. M., & González, M. C. (2015). Coupling human mobility and social ties. *Journal of The Royal Society Interface*, 12(105), 20141128.

Valenzuela, S., Park, N., & Kee, K. F. (2009). Is there social capital in a social network site?: Facebook use and college students' life satisfaction, trust, and participation. *Journal of computer-mediated communication*, 14(4), 875-901.

Tuxx. (z.d.). *Wegenkaart*. Geraadpleegd op 6 juni 2018, van <https://www.tuxx.nl/wegenkaart/>