

Smidig Mobilitet i Oslo

L2.1 State-of-the-art, Litteraturstudie

Versjon 1.0

Dato: 2014-12-18



Oslo kommune

Ruter#



Prosjektet er støttet av:



Revisjon

Versjon	Revidert av	Beskrivelse	Dato
0.1	Miriam Søgne Haugsbø	Utkast 1	juni 2013
0.2	Thomas Engen	Utkast 2	28.08.2014
1.0	Miriam Søgne Haugsbø	Endelig versjon	18.12.2014

Innhold	
REVISJON	II
INNHOOLD.....	III
1 INNLEDNING	1
2 BRUK AV GPS I REISEVANEUNDERSØKELSER.....	2
2.1 TRADISJONELLE REISEVANEUNDERSØKELSER.....	2
2.1.1 <i>Mer presise målinger av individuelle reisevaner</i>	3
2.2 TEKNOLOGI	3
2.2.1 <i>Mobiltelefoner og reisevaneundersøkelser</i>	4
2.2.2 <i>Algoritmer</i>	5
2.3 GPS-UTVALGET	6
2.3.1 <i>Responsrater</i>	6
2.3.2 <i>Brukeraksept og unit non-response</i>	7
2.3.3 <i>Brukerbelastning og item non-response</i>	8
2.3.4 <i>Insentiver</i>	9
2.3.5 <i>Personvern</i>	10
2.3.6 <i>GPS og systematiske målefeil</i>	11
2.4 DATA FRA EN GPS-BASERT UNDERSØKELSE.....	12
2.4.1 <i>GPS-baserte reisevaneundersøkelser</i>	12
2.5 LITTERATUR.....	13
3 TEKNIKKER FOR INNSAMLING OG ANALYSE AV DATA FRA HÅNDHOLDTE ENHETER	15
3.1 BAKGRUNN	15
3.2 INNSAMLING AV LOKALISERINGSDATA	16
3.2.1 <i>Utfordringer med smarttelefon</i>	16
3.2.2 <i>Mulighet for ikke-forstyrrende logging</i>	16
3.3 BEARBEIDING AV DATA	16
3.3.1 <i>Analysemetoder</i>	16
3.3.2 <i>Reisemønster</i>	16
3.3.3 <i>Utleddning av aktiviteter</i>	17
3.3.4 <i>Deteksjon av grupper</i>	18
3.4 UTLEDNING AV TRANSPORTMODI	18
3.4.1 <i>GPS-data</i>	18
3.4.2 <i>Bevegelsessensorer</i>	19
3.4.3 <i>GPS med supplerende data</i>	20
3.5 KONKLUSJON.....	20
3.6 LITTERATUR.....	20
4 OPPSUMMERING.....	22
4.1 ULIKE TEKNOLOGIER TIL ANVENDELSE – UTFORDRINGER OG MULIGHETER	22
4.2 KJENNETEGN VED UTVALGET I EN GPS-BASERT REISEVANEUNDERSØKELSE	23
VEDLEGG 1: SAMMENDRAG UTVALGTE STUDIER	25
SAMTIDIG DATAINNSAMLING OG OVERFØRING AV SANNTIDSFØRINGSINFORMASJON TIL BRUKER (BARBEAU M.FL. 2009)	25
GPS OG AUTOMATISK GENERERTE REISEVANEDATA (NITSCHKE M.FL. 2012)	26

PASSIV GPS-BASERT DATAINNSAMLING I ET KOMPLEKST URBANT MILJØ (CHEN M.FL. 2010)	27
SAMTIDIG PATH PREDICTION VED HJELP AV GPS OG MOBILTELEFONER (PERSAD-MAHARAJ M.FL. 2008).....	28
GENERATING PROBABILISTIC PATH OBSERVATION FROM GPS DATA FOR ROUTE CHOICE MODELING (CHEN M.FL. 2008)	28
BRUKERAKSEPT AV NY TEKNOLOGI I SPØRREUNDERSØKELSER (ROUX M.FL. 2009).....	28

1 Innledning

Dette notatet inneholder en gjennomgang av foreliggende studier innenfor GPS-baserte reisevaneundersøkelser, og er en sammenstilling av hovedkonklusjonene fra to litteraturstudier som er gjennomført innenfor prosjektet *Smidig Mobilitet i Oslo*; Haugsbø (2013) og Engen (2014).

Gjennomgangen av foreliggende studier og litteratur har i prosjektet dannet utgangspunktet for mange fruktbare diskusjoner og har gitt god input til teknologiutviklingen som er drevet innenfor prosjektet.

I kapittel 2 tas temaer som problemer med tradisjonelle reisevaneundersøkelser, og muligheter med GPS-baserte innsamlingsmetoder opp. Videre diskuteres ulike teknologiske løsninger og, samt at spesielle kjennetegn forbundet med IGPS-basert innsamling av data problematiseres. Kapitlet er basert på Haugsbø (2013). I kapittel 3 er fokus rettet mot teknologiske muligheter og utfordringer forbundet med GPS-baserte innsamlingsmetoder. Kapittel 3 er basert på Engen (2014). I kapittel 4 oppsummeres hovedkonklusjonene.

2 Bruk av GPS i reisevaneundersøkelser

GPS representerer en enorm mulighet i forbindelse med reisevaneundersøkelser da man gjennom å ta i bruk slike verktøy kan øke både kvaliteten og kvantiteten på data, samtidig som man minsker byrden for den enkelte deltager i undersøkelsen. Bruk av GPS-sporing vil generere nøyaktige data, og det vil være mulig å samle inn data for lange tidsperioder da brukerbelastningen er betydelig lavere enn ved tradisjonelle reisevaneundersøkelser (Chen m.fl. 2010).

Data som genereres fra tradisjonelle reisedagbøker er «bråkete» fordi man er avhengig av å rekruttere et representativt utvalg, samt at de individuelle rapporteringene er sannferdige og nøyaktige. Det er relativt greit å si noe om utvalgets representativitet, og det finnes velutviklede rutiner for å sikre dette hensynet, men samtidig ser man stadig voksende problemer knyttet til ikke-svar i tradisjonelle reisevaneundersøkelser. En større utfordring er likevel at det ikke er enkelt å si noe om i hvilken grad de rekrutterte deltagerne er sannferdige og ærlige i sine rapporteringer. Bruk av GPS og mobiltelefoner i reisevaneundersøkelser kan være med på å motvirke begge disse problemene (Gould 2013).

Når det kommer til bruk av GPS og mobiltelefoner i reisevaneundersøkelser representerer Norge og Oslo et godt case-studium. Oslo som by, og dens befolkning, skulle kunne forventes å gi gode data fra en slik studie. Denne er basert på det at Norge pr våren 2013 var et av seks land i verden hvor over 50 prosent av befolkningen eier en smart-telefon, og derfor inngår i utvalget til en slik undersøkelse (Gould 2013). Tall fra Statista viser at en av sju nordmenn har en smarttelefon per våren 2013¹. Videre er Ruter allerede tungt inne på dette teknologi-feltet, gjennom muligheten for rutesøk, sanntidsoppdateringer og billett kjøp på nett, og på mobiltelefonapplikasjoner. Schweigher (2011) framhever nettsiden som da het www.trafikanten.no, som et foregangseksempel på nettbaserte tjenester for kollektivtransportfeltet.

2.1 Tradisjonelle reisevaneundersøkelser

I reisevaneundersøkelsenes tidlige dager i 1950 og 1960-årene, var personlige intervjuer den rådende datainnsamlingsmetoden. Ulike fremgangsmåter ble brukt for å rekruttere respondenter, utvalgene var generelt store, opp til 30 000 respondenter, og undersøkelsene hadde generelt høge svarprosent. Relatert til kostnadsaspektet ble det på 1970-tallet mer og mer vanlig med rekruttering basert på telefonnummerregistre. Respondentene ble valgt ut ved hjelp av «random-digit dialling,» og det ble så sendt ut et spørreskjema i posten, som skulle returneres. En slik fremgangsmåte viste seg å generere lave svarprosent, og man gikk i løpet av 1990-tallet over til å samle inn dataene ved å ringe tilbake til respondentene. Utvalgene for reisevaneundersøkelsene ble samtidig mindre, og lå nå som regel på mellom to og tre tusen deltagere. Dette var motivert av økende kostnader, og av økende kunnskap om utvalgsstatistikk. Dagens reisevaneundersøkelser er basert på et variert utvalg av metoder for datainnsamling: ulike typer intervjuer og ofte i kombinasjon med ulike typer elektroniske og digitale hjelpemidler (Stopher 2009a).

Uavhengig av fremgangsmåte ser man et økende problem i forhold til en nedadgående svarprosent for slike undersøkelser. Dette er åpenbart knyttet til endring i fremgangsmåter for rekruttering, men det kan også knyttes til utviklingen av personlig markedsføring, og at folks generelle negative innstilling til telefonselgere har smittet over på telefonintervjuere (Stopher 2009a). Det er stor tro på at bruk av GPS-verktøy kan redusere denne trenden i mobilitetssurveyer (Gould 2013).

Det har lenge vært mulig å utføre GPS-baserte undersøkelser av reisevaner. Man har hovedsakelig basert seg på bruk av GPS-teknologi, og støttet seg på GSM i skyggeområder. På grunn av

¹ <http://www.dinside.no/921797/norge-paa-smartmobiltoppen>

batterikapasitetsproblemer, og størrelse/vekt på GPS-innretningen, har det vært problematisk å gjennomføre undersøkelser der enhetene er individer i kollektivnettet, eller som gående/syklende. Etter hvert har det blitt utviklet mindre og mer mobile innretninger, som det er mulig for individer å bære med seg. Videre har masseproduksjon gjort mobiltelefoner til allemannseie, og produktene på markedet blir stadig mer avanserte med tanke på posisjoneringstjenester. Potensialet knyttet til bruk av GPS i forbindelse med reisevaneundersøkelser er derfor stort (Barbeau m.fl. 2008).

2.1.1 Mer presise målinger av individuelle reisevaner

Bruk av GPS-sporing har mange fordeler i forbundet med nøyaktighet, og særlig i forhold til gange, sykkel og parkering. En annen fordel er at det er mulig å skaffe data for lengre tidsperioder enn den ene dagen som reisedagboken som regel begrenses til. GPS-data vil kunne gjøre rede for reisenes kompleksitet på en helt ny måte, og man kan oppnå en mye bredere forståelse av individenes handlinger under gitte forhold. Det blir ikke lenger nødvendig å spørre hva folk «ville ha gjort,» for eksempel dersom det oppstår kø på den ruten man vanligvis tar, da GPS-sporet vil kunne gjøre rede for individenes nøyaktige handlinger og bevegelser (Gould 2013).

Gange blir ofte underrapportert i reisevaneundersøkelser. Dette har sammenheng med at gåturer ofte er relativt korte, og at individene derfor enten glemmer å gjøre rede for turen, vurderer den som irrelevant på grunn av omfang, eller velger å ikke rapportere dem på grunn av at mange korte turer kan synes strevsomt å registrere. Videre er gange også en aktivitet i seg selv, som rekreasjon eller trening, og samtidig ikke helt klart skilt i fra gange som transport-metode, og det kan slik oppstå forvirring under rapportering (Gould 2013).

GPS-sporing kan forklare nøyaktig når, hvor og hvorfor individer går, og feilmarginen som må tas hensyn til går på teknisk utilstrekkelighet. Disse feilmarginene kan samtidig lett gjøres rede for, og den tilfører derfor ikke støy til datasettet i like stor grad som de usporbare feilkildene som reisedagboken ofte er belemret med. Sporingen kan videre gjøre rede for reisen i sammenheng med omgivelsene, gjennom at man får informasjon om hvilke ruter individene tar og ikke tar. Til slutt vil GPS-sporing tilføre verdifull informasjon om transitt-situasjoner og -tider. Fra gjennomførte reisevaneundersøkelser vet man at individene underrapporterer bytter, da de ikke har en fullstendig forståelse av reisekjeder, og at de samtidig overvurderer tiden som brukes på å vente i forbindelse med bytter (Gould 2013).

Sykkel er også gjerne underrapportert i tradisjonelle reisevaneundersøkelser, og mange av de ulike læringspotensialene som ble nevnt i forbindelse med gange, i forhold til nøyaktighet og veivalg, gjelder også her. Skillet mellom reiser som er rene transportetapper, og reiser som er forbundet med rekreasjon, er gjerne enda vanskeligere å trekke når det kommer til sykkel. Et interessant moment i denne forbindelse er at syklister gjerne er unge menn, som samtidig er en gruppe som gjerne unngår tradisjonelle surveyer. En GPS-sporing derimot, kan tenkes å være spesielt aktuelt, da man ser for seg at denne gruppen mennesker også er relativt teknologi-interesserte (Murakami 2008, i Gould 2013).

I forhold til parkering kan GPS-sporing lære oss mye, da de fleste transportmodeller ikke i stor nok grad tar hensyn til tidsbruk og kostnader forbundet med parkering. En GPS-sporing vil gjøre rede for hele reisen med bil, inkludert tid brukt til parkering, på en objektiv måte. Sporingen kan videre kartlegge om organiseringen av parkering i seg selv genererer «kalde starter,» forbundet med at individene flytter bilen til en ny parkeringsplass etter hvert som parkeringstiden på et sted går ut (Gould 2013).

2.2 Teknologi

Det er tre ulike veiskiller i forhold til hvilken teknologi man skal velge for å gjennomføre en studie som omfatter sporing av reiser ved hjelp av GPS eller tilsvarende:

- Kjøretøy- eller personsporing
- Interaktiv- eller passiv sporing
- Logging eller samtidig overføring av data

Studiene som sporer reisevaner har som regel hatt kjøretøy som enheter, dette har sammenheng med batterikapasitet og vekt/størrelse på enhetene. Etter hvert har det blitt utviklet mindre og mer anvendelige enheter til slike studier, som individene har mulighet til å bære med seg. Det har likevel ikke vært gjennomført særlig mange full-skala GPS-baserte studier med individer som enheter (Stopher 2009b). Dette har sammenheng med at en studie med individer som enheter innebærer sporing i et mer finmasket nettverk; kollektivnettverket i et byområde, til sammenligning med det generelle veinettet. Videre har man problemer relatert til såkalte «urbane dalfører,» som oppstår når individene beveger seg mellom høye bygninger, og i tillegg vil man ha problemer med å skille ulike lokaliteter fra hverandre, relatert til variert arealbruk og tett befolkede områder. Korte og hyppige reiser, og variert reisemiddelbruk, gjør det utfordrende å registrere alle reisens detaljer (Chen m.fl. 2010).

Et annet kjennetegn for de hittil gjennomførte studiene, er at de har vært basert på interaktiv sporing. Deltageren blir bedt om å skru på enheten når reisen starter, og om å oppgi visse typer informasjon i etterkant av reisen, gjennom kommunikasjon på e-post/telefon. Studiene som er gjennomført krever ulik grad av medvirkning fra respondentene, og er derfor interaktive i varierende grad. Sporing av individene ved hjelp av ulike typer GPS har ofte vært praktisert som et supplement til den tradisjonelle reisedagboken, noe som innebærer stor grad av interaksjon med respondenten. Videre er det i mange GPS-baserte studier lagt inn en reisedagbok-komponent, i den forstand at man gjerne spores tilnærmet passivt over noen dager, mens man fyller ut reisedagbok for en av dagene. Dette for å validere sporingen mot reisedagboken. Eksempler på en slik fremgangsmåte finner man i Kohla og Meschvik (2013) og Stopher m.fl. (2013).

Passiv sporing har vært lite utbredt, hovedsakelig på grunn av manglende tilgjengelig teknologi, men valget mellom interaktiv eller passiv sporing får også følger for personvern og brukerssept. Den teknologiske utviklingen kan på sikt gjøre brukermedvirkning potensielt overflødig. Gjennom bruk av algoritmer kan man avdekke ulike karaktertrekk ved reisen, f.eks. reisemiddel, reiserute osv., og brukerbelastningen blir redusert ytterligere ved reisevaneundersøkelser (Stopher 2009b). Eksempler på studier som går langt i bruk av algoritmer og gjennomfører tilnærmet passiv sporing er Persad-Maharaj m.fl. (2008), Barbeau m.fl. (2008), Chen m.fl. (2010). Det å ta i bruk passiv sporing vil alltid begrenses av personvern hensyn.

Et tredje kjennetegn ved GPS-baserte reisevaneundersøkelser er at de stort sett har basert seg på en logging av data, ikke en samtidig overføring. Dette har å gjøre med at det ikke har vært nødvendig i forhold til problemstillinger som har vært undersøkt, og videre at det er knyttet utfordringer til f.eks. batterikapasitet i forhold til samtidig overføring (Stopher 2009b). Også her kommer personvern hensyn inn i bildet, da samtidig overføring av data regnes som en større inngripen i privatlivet enn logging av data. Eksempler på logging av datapunkter ved sporing er Stopher (2013) og Roux m.fl. (2008), mens Stenneth (2011) og Barbeau m.fl. (2008) baserer seg på samtidig overføring.

2.2.1 Mobiltelefoner og reisevaneundersøkelser

Mobiltelefoner har så lenge teknologien har vært tilgjengelig for folk flest, hatt en viktig rolle i forbindelse med mobilitet og reiseaktivitet. I første omgang gjorde mobiltelefoner det mulig å koordinere møtesteder med andre personer under reisen, og etter hvert som mobiltelefonene har blitt utstyrt med internett, har man et nyttig verktøy for ruteplanlegging alltid tilgjengelig. Bruk av mobiltelefoner som datainnsamlingsverktøy i reisevaneundersøkelser er derfor nærliggende med tanke på denne teknologiens gjeldende nedslagsfelt (Gould 2013).

Fremtidens utfordringer knyttet til reisevaneundersøkelser handler derfor om tilpasning til fremtidens mobiltelefoner. Gould (2013) gjør rede for to ulike tilnærminger til det å ta i bruk mobiltelefoner i reisevaneundersøkelser. Et mulig bruksområde for mobiltelefoner er gjennom passiv sporing basert på mobiltelefonens applikasjoner. Fordelen med enkeltstående GPS-enheter i forhold til mobiltelefoner er at de er mer nøyaktige. Ulempene er relatert til det at en egen «dings» forbundet med studien, gjør deltagerne oppmerksomme på at de blir sporet. Mobiltelefonen derimot er i dagens samfunn en naturlig følgesvenn på selv den korteste reisen og kan derfor ikke tenkes påvirke individene i særlig grad. Videre blir det fremhevet at mobiltelefonindustrien er i ekstremt rask utvikling, og man kan derfor forvente store fremskritt innenfor dette området i den nærmeste tiden. Potensialet i forhold til reisevaneundersøkelser vil derfor etter hvert bli enda større enn hva man er klar over i dag (Gould 2013).

2.2.2 Algoritmer

Det er et mangfold av ulike algoritmer i bruk til å avdekke ulike aspekter ved de ulike reisene. Et interessant moment som blir fremhevet av Gould (2013) er at utviklingen av algoritmer til bruk under slik datainnsamling, har mye til felles med utviklingen av survey-spørsmål i tradisjonelle reisevaneundersøkelser. Evnen til å utnytte det potensialet som GPS-sporing innebærer, er avhengig av gode algoritmer som gjør rede for de ulike aspektene ved reisen på en god måte, samtidig som man er bevisst hva de ulike algoritmene eventuelt kan «skjule» av kunnskap. Flere ulike studier avdekker ulike detaljer ved reisene ved hjelp av algoritmer, for eksempel reisemiddel (Stopher m.fl. 2013; Stenneth 2011; Chen m.fl. 2011), reisemål (Stopher m.fl. 2013; Chen m.fl. 2011) og reiserute (Persad-Maharaj 2008; Chen m.fl. 2011).

Chen m.fl. (2011) gjør rede for bruk av algoritmer til å avdekke reisenes detaljer på en god måte. Denne studien tester en rekke GIS-algoritmer, som automatisk avdekker reisemiddel og antyder målsettingen med reisen. Prosessen er delt inn i fire ulike deler. Første skritt er innsamling av data til to ulike datasett. I det første datasettet inngikk 25 tilsatte ved the New York Metropolitan Transportation Council, som ble bedt om å bære med seg en GPS-enhet gjennom en arbeidsdag. Det andre datasettet bestod av reisevanedata fra 24 studenter ved the City University of New York, som ble bedt om å bære med seg senderen over en periode på fem virkedager.

Å gjennomføre GPS-baserte undersøkelser i New York er utfordrende hovedsakelig på grunn av urbane dalfører, et ekstremt tett befolket område med et veldig variert arealbruksmønster, og komplekse transport-nettverk. Man har store områder som er utenfor GPS-dekning, og reisene som foretas er preget av geografien: korte og hyppige reiser og variert reisemiddelbruk, som gjør det utfordrende å registrere alle reisens detaljer.

Enheter logget hver persons posisjon hvert femte sekund, inkludert; dato, tid, lengde- og breddegrad, fart osv. Hver respondent skrudde på GPS-enheter hver morgen, og skrudde den av hver kveld når man ikke skulle ut igjen. Hver enkelt ble videre bedt om å fylle ut en detaljert reisedagbok (navn på destinasjon, nøyaktig adresse til destinasjon eller skjæringspunkt, reisemiddel, avgang- og ankomsttider, målsetting med reisen) for en dag i løpet av studien. I tillegg spurte man etter eventuelt eierskap til bil og om å oppgi steder man besøkte ofte. Poenget med å be informantene om å fylle ut en reisedagbok i tillegg til GPS-sporingen var å skulle kunne sammenligne disse to informasjonskildene.

Andre steg i denne studien var å utarbeide ulike kart-lag, som representerte ulike transportmidler, f.eks. veinettet, bussruter inkludert stopp, undergrunnsystemet inkludert innganger/utganger og pendlertogruter og -stopp. De ulike transportmidlene er som sagt organisert i ulike lag, men mange plasser vil de ulike lagene krysse hverandre i kartet, uten nødvendigvis å være knyttet sammen. Undergrunnsbanens nettverk har gjerne kryssende, eller delvis sammenfallende plassering i forhold til

bussnettet, som ligger som et lag oppå. Videre er det mulig at en bilreise foretas i mer eller mindre samme rutetrase som bussen. Reisen blir foretatt på samme geografiske punkt, men det er likevel ikke nødvendigvis kontakt. F.eks. en respondent som blir knyttet til et punkt som er identifisert til en undergrunns-nedgang, vil være knyttet til undergrunns-nettet helt til personen blir knyttet til et punkt hvor det er mulig å komme seg ut av undergrunns-nettet igjen. Gjennom å kombinere informasjonen fra de ulike kartlagene kan man utarbeide en database som på en enkel og grei måte kan nyttes til å identifisere reisemiddel og reisemål.

Det tredje steget i undersøkelsen er myntet på å identifisere reisemiddel for den enkelte reisen. Dette gjøres ved å først identifisere ulike segmenter av reisen, basert på informasjonen fra GPS-punktene som er samlet inn. Gjennom å sjekke om det er hull i GPS-overføringene (et hull blir regnet som 120 sekunder eller 250 meter), kan man skille mellom undergrunns-transport og transport på gateplan, da hull betyr tap av GPS signal, som taler for undergrunn. Neste steg er å identifisere aktuelle punkter hvor det er sannsynlig at bruker endret reisemiddel. Dette blir gjort gjennom at man antar at enhver reisemiddelendring er forbundet med en kort gåtur, som i denne studien blir definert som over et minutt. Etter at segmentene i reisen er identifisert forsøker man å knytte hvert segment til et reisemiddel, gjennom en algoritme basert på de ulike reisemidlenes kjennetegn.

Det fjerde steget identifiserer målsettingen med reisen. Ut i fra opprinnelsessteder og reisemål kan en arealbruksalgoritme koble destinasjonen med de aktivitetene som man vet foregår på dette geografiske punktet. Der slike automatiske koblinger ikke er mulig kan man sannsynliggjøre at brukeren hadde en gitt målsetting med turen ved å basere seg på tidspunkt, brukerens historiske reisevaner og arealbruk i området rundt punktet.

Stenneth (2011) går enda lenger i retning passiv sporing. Her kombinerer man informasjon fra GPS sensorer på mobiltelefoner, med informasjon om det underliggende transportnettverket, inkludert sanntidsinformasjon om lokalisering av for eksempel busser i nettet, til å identifisere reisemiddel. I dette prosjektet var man i stand til å skille mellom ulike typer motorisert ferdsel, noe som har vært problematisk frem til nå. En interessant nyvinning i dette prosjektet er algoritmer som identifiserer «average bus closeness», «average rail closeness» og «candidate bus closeness.»

Undersøkelsene er basert på Chicago sitt sanntidsinformasjonssystem for kollektivnettet. Alle busser som hører til Chicago Transit Authority, som er det største busselskapet, er utstyrt med GPS-sensorer, og spores kontinuerlig i systemet. Informasjonen som er tilgjengelig inkluderer rute, lengdegrad, breddegrad, endestopp, bus ID og retning. En indeks basert på alle postnumre i området, er koblet sammen med en liste over alternative busstopp, bussruter og ruter for skinnegående, som hører til hvert enkelt postnummer. Når en bruker blir registrert med et geopunkt som hører til et gitt postnummer er det på denne måten en enkel sak å identifisere ulike alternative reisemiddel og ruter.

2.3 GPS-utvalget

Mobilitetssurveyer må være representative i forhold til tid, rom og sosiodemografiske faktorer, i tillegg til representativitet over ulike transportmiddel og mellom ulike typer reisemål. Å begrense undersøkelsens univers til den delen av befolkningen som har en smarttelefon, fører til at man må forholde seg til nye utfordringer knyttet til representativitet (Nitsche et. al 2012). Det er derfor viktig å undersøke kjennetegn ved utvalg til GPS-baserte studier, for slik å kunne utvikle strategier og verktøy for å sikre representativiteten i fremtidige GPS-utvalg (Bricka 2009).

2.3.1 Responsrater

Bricka (2009) gjennomfører en studie som tar for seg responsrater i GPS-baserte reisevaneundersøkelser. Man ser på utvalgene fra to ulike gjennomførte studier, og undersøker i hvilken

grad det er avvik mellom utvalgene fra hovedundersøkelsen og GPS-undersøkelsen. Både i forhold til hvilke grupper som sa seg villige til å være med på den GPS-baserte studien (unit non-response), og hvilke grupper som faktisk gjennomførte GPS-studien på en tilfredsstillende måte (item non-response). Undersøkelsen støtter seg på en gjennomgang av litteratur som omhandler rekruttering og representativitet i tradisjonelle reisevaneundersøkelser, og man trekker i tillegg inn litteratur fra teori som tar for seg spredning av nye teknologier, fra markedsforskning og informasjonsteknologilitteratur. Den underliggende hypotesen for studien er at:

[N]on-response in GPS-based travel surveys is a function of the non-response associated with traditional travel surveys and the characteristics of those that do not quickly embrace new technologies (Bricka 2009:409).

I tradisjonelle reisevaneundersøkelser har svarprosenten sammenheng med både demografi og reisevanekarakteristika hos respondentene. De vanligste demografiske karakteristika for non-respons i utvalgene i tradisjonelle reisevaneundersøkelser er lav inntekt, lav alder (under 25 år), lavere utdanning, medlem av en husholdning med mange medlemmer, minoritetsbakgrunn og bosted i et urbant område. Videre hadde de å ha under gjennomsnittet antall reiser, og det å ikke bruke offentlig transport, innvirkning i negativ retning i forhold til om individet ble med i utvalget.

I en GPS-basert studie kan man forvente at individenes forhold til teknologi vil påvirke i hvilken grad de er villige til å delta som informanter i studien eller ikke. Det er vanlig å anta at aksept av ny teknologi avhenger av et individs personlige opplevde nytte av den gitte teknologien, og brukervennligheten til produktet. Teorien tilsier at aksept av teknologi generelt er høyere blant menn, ungdom, de med høyere utdanning og de med høyere inntekt (Bricka 2009).

Et utvalg for en GPS-basert reisevaneundersøkelse vil derfor være en funksjon av individenes demografiske karakteristika, deres reisevaneprofiler og deres aksept av teknologi. Til sammenligning med et utvalg til en tradisjonell reisedagboks-undersøkelse kan en GPS-undersøkelse derfor forventes å ha enda færre respondenter fra grupper med lav utdanning og lav inntekt, men at representativiteten for de yngre delene av befolkningen vil bli bedre (Bricka 2009).

2.3.2 Brukeraksept og unit non-response

Den hittil mest utbredte bruken av GPS-sporing har vært som et tillegg til den tradisjonelle reisedagboken, for å gi et bilde på validiteten til resultatene av denne. Det finnes bare et fåtall rene GPS-studier f.eks. Stopher m.fl. (2013). Når GPS blir brukt som et supplement til tradisjonelle reisevaneundersøkelser, er en svært utbredt tilnærming at respondenten først går igjennom det vanlige intervjuet, før spørsmålet om å delta i GPS-sporingen blir tatt opp helt på slutten av samtalen. Dette er gjort for å hindre at et slikt spørsmål skal ha innvirkning på deltagelsen i den tradisjonelle delen av studien. Som en effekt av dette vil respondentenes villighet til å delta i GPS-sporingen være påvirket av byrden som dette intervjuet representerer (Stopher 2009b).

I slike tilfeller, der GPS-utvalget er basert på et spørsmål om villighet til deltagelse stilt til hovedutvalget, har man sett at om lag en tredjedel av respondentene sier seg villig til å delta i en slik studie. Det blir hevdet at svarprosenten er betydelig avhengig av det at respondenten allerede har gjennomgått et omfattende intervju, og man ser for seg at responsrater for en studie som utelukkende er basert på GPS-sporing, vil kunne oppnå samme prosenter som de tradisjonelle reisedagbøkene (Stopher 2009b).

Bricka (2009) tar for seg to studier fra USA, en i Kansas og en i Washington, DC, der man har gjennomført reisevaneundersøkelser basert på en tradisjonell reisedagbok, og hvor man også har gjennomført en tilleggs-studie, der deler av utvalget har blitt sporet ved hjelp av GPS. De to studiene skiller seg fra hverandre i forhold til hvordan GPS-utvalget blir trukket ut fra hovedutvalget.

Kansas City Household Travel Survey ble gjennomført våren 2004. Data om demografi og reisevaner ble samlet inn for 3049 tilfeldig utvalgte husholdninger, og tillegg ble 294 husholdninger utstyrt med en GPS logger. Studien praktiserte en «opt in- approach,» der husholdningene fikk tilbudet om å gjennomføre en GPS-sporing dersom de hadde interesse av dette. Washington, DC/Baltimore Regional Household Travel Survey fra januar 2008 samlet inn tradisjonelle reisevanedata for 10 836 husholdninger, fra et utvalg på 15 000. Av disse var 1600 husholdninger inkludert i et GPS-underutvalg, og 850 deltok. Husholdningene i GPS-utvalget var tilfeldig utvalgt, og respondentene ble informert om at de var trukket ut til denne delen av studien, men at de kunne reservere seg mot dette dersom de ønsket det. Husholdninger som var trukket ut fikk muligheten til å «opt out» men andre husholdninger fikk ikke muligheten til å «opt in.»

Det er interessant å se hvilke konsekvenser de to rekrutteringsmetodene fikk for utvalgets sammensetning. Sammenlignet med hovedutvalget fra undersøkelsen fra Kansas, har GPS-utvalget totalt større underrepresentasjon for grupper med lavere utdanning, lavere inntekt og høyere alder. Dette er i tråd med teorien for teknologiaksept. Forskjellene var signifikant forskjellige i forhold til størrelse på husholdningen, geografi, inntekt, minoritetsbakgrunn og alder, men ikke for utdanning. For Washington, DC/Baltimore undersøkelsen, viste en sammenligning av de to utvalgene at respondenter fra større husholdninger og yngre respondenter, var bedre representerte i GPS-utvalget, mens lav-inntekt husholdninger og respondenter med minoritetsbakgrunn var i større grad underrepresenterte. Også i dette tilfellet var funnene konsistente med teknologiaksept-litteraturen. Utvalgene var signifikant forskjellige i forhold til husholdningsstørrelse, inntekt, minoritetsbakgrunn og alder.

Oppsummert indikerer dette at non-respons problemene som er relevante for tradisjonelle reisevaneundersøkelser, også gjelder for GPS-baserte studier, og at enkelte av de tradisjonelle problemene blir større med et slikt utvalg. Begge studiene som Bricka (2009) gjør rede for finner at yngre respondenter i større grad er villige til å la seg spore. Det er derfor interessant å merke seg at Sneade (2013) på den andre siden finner at unge mennesker var mindre villige til å bære med seg en GPS-enhet.

2.3.3 Brukerbelastning og item non-response

Bruk av GPS for innsamling av data forventes å minske belastningen på respondenten, og det forventes derfor at item non-response for undersøkelsen minskes tilsvarende. Item non-response er et mål på hvor nøyaktig respondentene har oppgitt dataene. Denne antagelsen blir imidlertid satt spørsmålsteget ved av Bricka (2009). Dette har sammenheng med at GPS-studier hittil ikke har vært gjennomført som et mål i seg selv, men som et middel for å skulle kunne forbedre de tradisjonelle reisedagbøkene. GPS-sporingen kommer i slike tilfeller som et tillegg, ikke som en erstatning til de tradisjonelle reisevaneundersøkelsene, og brukerbelastningen kan derfor ikke sies å ha minsket. Det har hittil vært relativt lite fokus på brukerbelastning under GPS-basert datainnsamling, og gjennomgående drøftinger av dette temaet er derfor ikke tilgjengelig i litteraturen.

Bricka (2009) finner at dagens bruk av GPS i reisevaneundersøkelser tvert imot fører til en større belastning på brukerne. Reisevanedata blir i dag samlet inn på flere ulike måter, og denne studien kvantifiserer og teller opp brukerbelastningen som de ulike metodene, og kombinasjonene av disse, betyr for respondentene. Designet som gir høyest brukerbelastning er der hvor respondentene gjennomfører en tradisjonell reisedagboks-undersøkelse, som blir supplert med GPS-data for valideringshensyn, og som igjen utgjør basis for en PR-samtale der respondenten skal gjøre ytterligere rede for reisene sine. Et slikt design betyr en fordobling av brukerbelastningen, sammenlignet med en studie som benytter seg av den tradisjonelle reisevanedagboka.

Stopher (2009b) undersøker responsrater i ulike GPS-studier, og finner at det er generelt lite fokus på disse tallene i de ulike studiene, men at det ser ut som om at tallene for item non-response i GPS-baserte og tradisjonelle reisevaneundersøkelser er om lag like.

Sameksistens med den daglige bruken av mobiltelefonen er et viktig aspekt i forbindelse med item non-response. Dersom respondenten opplever at mobiltelefonen ikke fungerer på en tilfredsstillende måte mens applikasjonen kjører, kan dette føre til at man ikke logger alle reisene sine. De mest aktuelle problemstillingene er at programmet kan gjøre mobilens andre funksjoner mindre smidige, eller at mobiltelefonens batteri kan tappes som følge av at applikasjonen kjører. Det er derfor helt nødvendig å satse på utvikling av såkalt *smart resource management* for disse applikasjonene. Eksempelvis gjennom det å kalkulere så få GPS-posisjoner som er absolutt nødvendig, og deretter å bare overføre informasjon som er kritisk for å rekonstruere brukeren sin reiserute (Barbeau et. al 2008).

Et annet hinder for kan være det at deltagelse i undersøkelsen fører til merarbeid for den enkelte, gjennom at man må skru på applikasjonen før hver reise, og gjøre rede for f.eks. reisemål og -middel m.m. Passive og automatiske versjoner av slike applikasjoner, som ikke krever manuell input fra brukeren er under utvikling. Slike applikasjoner må klare å finne nok informasjon fra GPS-data og fra omgivelsene, til å kunne avdekke opplysninger som i dag blir samlet inn gjennom samhandling med brukeren. Dette kan gjøres gjennom algoritmer der slik informasjon blir forsøkt avdekket, basert på for eksempel kart som omfatter vegnett, kollektivtrafikk og ulike tjenester (Barbeau et. al 2008). Det er også gjort vellykkede forsøk med akselerasjonssensorer til å identifisere reisemiddel (Nitsche et. al 2012). Her kommer hensyn til personvernet inn i bildet igjen.

Praktiske utfordringer som kan føre til item-non response inkluderer også det at det ikke er alle grupper som har like gode forutsetninger for å ta i bruk teknologien som kreves for å delta i slike undersøkelser. Det blir derfor lagt vekt på å gjøre brukergrensesnittet til en mobilapplikasjon så enkelt og intuitivt som mulig. Videre er det viktig at en mobilapplikasjon som ønsker å nå ut til en stor gruppe mennesker kan brukes på mer enn en type maskinvare eller programvare uten at man må gjøre store endringer i kodingen av applikasjonen sin programvare (Nitsche et. al 2012).

2.3.4 Insentiver

Rekrutteringen til slike studier kan økes gjennom ulike former for insentiver. Disse tiltakene bør tilpasses for å sikre rekruttering fra de gruppene i befolkningen som har mindre sannsynlighet for å delta. I denne sammenheng er det viktig å huske at aksept og reaksjon på insentiver er kulturelt betinget, og vil variere mellom land (Stopher m.fl. 2009).

Ulike studier diskuterer flere forskjellige insentiver for å sikre deltagelse i GPS-sporing. For eksempel ble det drøftet å betale for mobildata-kostnader for respondenten i perioden som undersøkelsen foregår, for å sikre at respondenten ikke har utgifter forbundet med sporingen. Videre er det foreslått å levere samtidige trafikkoppdateringer til brukeren langs ulike predikerte ruter. Ved å basere seg på tidligere reiser som brukeren har gjennomført kan man forutse alternative reiseruter, og advare brukeren via tekstmeldinger eller lignende dersom hendelser blir oppdaget (Barbeau m.fl.2008). Denne tanken blir videreutviklet, og man ser for seg at det er interessant å tilby deltagerne i en slik studie personlig tilpasset reklame som er relevant for området som han eller hun er i ferd med å bevege seg inn i (Stenneth 2011, Persad-Maharaj m.fl. 2008). Den ytterste konsekvens av denne tankegangen er å oversende reklame/kuponger for bensin, dersom reisemiddel-algoritmen avdekker at reisen foregår med bil (Stenneth 2011).

Et annet insentiv som er diskutert er å gi deltagerne i en slik studie tilbud om informasjon relatert til reisene de har gjennomført, for eksempel karbonavtrykk, dersom reisen foregikk med motorisert

kjøretøy, eller antall kalorier forbrent, dersom reisen foregikk ved hjelp av egen maskin (Stenneth 2011). Til sist er det mulig å kompensere deltagerne i en slik studie økonomisk, enten alle deltagerne, eller bare grupper med lav deltagelsesgrad (Stopher m.fl. 2013, Sneade 2013).

2.3.5 Personvern

Personvernet er et viktig hensyn i forhold til gjennomføring av surveyer, og kanskje er dette ekstra ømfintlig i forhold til reisevaneundersøkelser. Kunnskap om reiser gir potensielt innsikt i sensitive personlige data, for eksempel politisk og religiøst syn, helsetilstand, bekjenskaper og så videre. I en tradisjonell reisevaneundersøkelse har individet fullstendig kontroll over de data de oppgir i studien, gjennom at data er basert utelukkende på respondentens egen redegjørelse, og hensynet til personvern er ikke truet på samme måte som ved gjennomføringen av en digital sporing av respondentens mobiltelefon. En sporing av individet innebærer at også reiser som individet regner som private, og som de kanskje ikke ville ha oppgitt under en tradisjonell reisevaneundersøkelse, blir logget og utgjør en del av datagrunnlaget. Individet mister på denne måten delvis kontrollen over hva man deler, og personvernet er derfor i slike undersøkelser i enda større grad et viktig tema (Smoreda m.fl. 2013).

De ulike fremgangsmåtene for sporing innebærer ulike følger for personvernet. Sporing av kjøretøy er mindre inngripende enn sporing av bevegelser på personnivå, og logging av data er mindre inngripende enn samtidig overføring. Det viktigste teknologiske valget, i forhold til personvernet, er valget mellom aktiv og passiv sporing. En aktiv sporing innebærer at individet selv setter i gang sporingen, når en reise tar til. Slik beholder individet kontroll over hvilke reiser som blir sporet. Ved passiv sporing er sporingen kontinuerlig, og det er ikke nødvendig at individet starter applikasjonen før en reise. Individet har dermed ikke lenger kontrollen over hvilken informasjon som blir delt (Stopher 2009b).

Det er ulike tiltak som kan settes i verk, for å bedre personvernet for deltagerne i en studie som sporer deltagerne digitalt. Det vil hele tiden være en avveining mellom datakvalitet og personvern. Barbeau m.fl. (2008) foreslår oppretting av sensitive soner/tidsperioder hvor GPS-sporing ikke skal foregå, for å sikre bedre personvern for den enkelte deltager.

Stenneth og Yu (2011) tar for seg tredjeparters interesse for individers søk og oppdateringer på nettet, kombinert med kunnskap om lokaliteter. Det er interessant for kommersielle aktører at et individ er på en gitt lokalitet, og søker opp informasjon om et gitt forhold på nettet. Det å knytte denne informasjonen sammen er nøkkelen. Gjennom en «K-anonymity model» kan man anonymisere individene gjennom å sikre at de ikke kan skilles fra et gitt antall andre individer som er i området. De ulike individene blir i neste omgang gruppert sammen i forhold til hvilket reisemiddel de bruker, for å sikre at de ikke kan identifiseres på denne måten.

Det er interessant å undersøke i hvilken grad personvern spiller inn i forhold til individers vilje til å delta i en undersøkelse som tar sikte på å spore reisene deres. Kohla og Meschik (2013) sammenligner den metodiske fremgangsmåten som den tradisjonelle reisedagboken representerer med en GPS-basert datainnsamling, og bruker data fra den østerrikske mobilitetsundersøkelsen. Her finner man at bare 1,9 % av respondentene som motsatte seg sporing ved hjelp av GPS, gjorde dette ut i fra personvernsomsyn. De vanligste grunnene til avvísning av GPS-sporing var det at respondentene opplevde sin egen reisehverdag som lite representativ eller lite interessant å studere. Videre viser Swann og Stopher (2008), gjennom en fokusgruppestudie av respondenter fra en GPS-basert undersøkelse, at respondentene ikke hadde følt at personvernet deres var truet under gjennomføringen av studien.

2.3.6 GPS og systematiske målefeil

Det er knyttet store forventninger til bruk av GPS i reisevaneundersøkelser, men samtidig vet man lite om slike tilnærmingers potensielle systematiske målefeil. Swann og Stopher (2008) gjennomfører fokusgrupper med respondenter fra en GPS-basert studie, med siktemål om å kartlegge GPS-springens innvirkning på individuelle handlinger, samt respondentenes personlige erfaringer med å delta i en slik undersøkelse. Fokus var på å kartlegge om det var spesielle segment av befolkningen som hadde problemer med å bruke enheten, og om det var spesielle typer reiser som tenderte til å bli underrapportert.

Fire ulike fokusgrupper ble rekrutterte blant respondenter fra en GPS-basert undersøkelse gjennomført i Adelaide, Australia. Fokusgruppene ble designet med det formål å skulle kunne kontrastere forskjellene mellom «good performers» og «bad performers» blant respondentene. Om respondentene havnet i den ene eller den andre gruppen ble basert på hvor mange dager med GPS-data de hadde bidratt med i studien. Gruppene hadde ellers en variert demografisk miks i forhold til kjønn, familiestruktur og yrkesaktivitet. Det ble lagt mye arbeid ned i rekruttering og oppfølging, samt at man tilbød deltagerne en økonomisk kompensasjon.

Respondentene rapporterte i stor grad at de hadde forstått oppgaven de hadde blitt bedt om å utføre, men enkelte «bad performers» hadde misforstått hvilke aktiviteter som skulle logges. I denne gruppen med «bad performers», fantes det også respondenter som oppgav at de hadde hatt problemer med å forstå hvordan enheten fungerte. Respondentenes forståelse av innretningens funksjoner syntes imidlertid å være mer avhengig av deres tidligere erfaringer med slike teknologier, heller enn innføringen som var blitt gitt i forhold til denne spesifikke studien.

Respondentene var generelt veldig motiverte for å sikre gode data. De hadde etter beste evne ladet enheten, og forsøkt å sikre at den fikk signal til å måle reisen til enhver tid. Enkelte var bekymret for at deres gjennomføring av studien hadde generert data som ikke gjorde rede for bevegelsene på en god nok måte, og at dette igjen kunne føre til et dårligere fremtidig kollektivtilbud dersom man la denne studien til grunn. Et problem som ble rapportert var at det kunne være vanskelig å huske å ta med seg enheten i en ellers travel hverdag, og at enheten bla utladet før alle dagens reiser var gjennomførte. Ingen respondenter oppgav at de hadde lagt igjen enheten hjemme med vilje i løpet av studien.

Det var ingen av respondentene som følte at mengden papirarbeid som var forbundet med studien hadde vært «for mye,» og det kom frem at enkelte respondenter mente de kunne ha bidratt med mer informasjon. Fokusgruppene virket videre å fungere som en slags forsikring for respondentene, om at deres bidrag hadde blitt satt pris på. Det var særlig tydelig at mange satte pris på å få sjansen til å komme med personlige innspill på forhold som gjaldt transportområdet. Deltagernes engasjement tyder på at de som sier seg villige til å delta i slike undersøkelser gjerne har en interesse for transportområdet som går ut over den jevne befolkningen.

Et interessant funn fra fokusgruppene var at ingen av respondentene mente at deltagelsen i studien hadde invadert deres personvern, men tvert imot, at dette problemet var overdrevet i forhold til transportområdet. Dette er et uttrykk for et samfunn med en generelt høy grad av sosial tillit, men kan også knyttes til sammensetningen av fokusgruppene. Både i forhold til at deltagerne gjerne hadde et særlig engasjement forbundet med transportområdet, og videre til at de som sier seg villige til å delta i en GPS-basert undersøkelse ikke er de mest konservative når det kommer til datalagring og personvern.

Denne studien kan ikke si noe om GPS er en bedre datainnsamlingsmetode sammenlignet med tradisjonelle reisedagbøker eller ikke, men sier noe om kvaliteten på data fra GPS-spring isolert sett. Fokusgruppene avdekket ingen systematiske målefeil, men viste heller at respondentene hadde vært svært pliktoppfyllende i forhold til å forsøke å sikre tilfredsstillende data.

2.4 Data fra en GPS-basert undersøkelse

Hovedfordelen med data fra en GPS-sporing er at de er mer nøyaktige, både i forhold til posisjoner, tid og avstander, sammenlignet med tradisjonelle datainnsamlingsmetoder. Videre gjør den lave belastningen på respondentene til at det er mulig å samle inn data for lengre tidsperioder enn tidligere. På denne måten er både kvaliteten og kvantiteten på data bedret. På den andre siden er det en del interessante faktorer som i utgangspunktet ikke kan gjøres rede for gjennom GPS-sporing: reisemiddel, størrelse på reisefølget, målsetting med reisen, kostnader forbundet med reisen, alternativer til reisen som ble gjennomført og inntrykk/holdninger/preferanser hos respondenten (Stopher 2009b). Gjennom bruk av algoritmer kan man i dag avdekke reisemiddel og målsetting for de fleste reiser, som for eksempel Chen m.fl. (2010) har vist, men samtidig er informasjon om kostnader og størrelse på reisefølget er svært interessante reisedetaljer som man ikke kan passivt kartlegge ved hjelp av GPS-verktøy (Stopher 2009b).

Data som er basert på en GPS-sporing har andre karakteristika enn data fra en tradisjonell undersøkelse, og målsettingen om å erstatte de tradisjonelle reisedagbøkene, må ta dette med i betraktningen. En egnet, og mye brukt fremgangsmåte, for å sikre at man har et realistisk bilde av hva de GPS-baserte dataene forteller, og i hvilken grad de er overlegne tidligere data, er å samle inn både tradisjonelle og GPS-baserte data for en gruppe respondenter, og se i hvilken grad de to datasettene beskriver et noen lunde likt bilde.

Chen m.fl. (2011) gjennomfører en undersøkelse i New York, der man blant annet har som målsetting å sammenligne GPS-data og tradisjonelle data. Enheten som ble brukt i denne studien logget hver persons posisjon hvert femte sekund, inkludert; dato, tid, lengde- og breddegrad, fart osv. Hver respondent skrudde på GPS-enheten hver morgen, og skrudde den av hver kveld når man ikke skulle ut igjen. Hver enkelt ble videre bedt om å fylle ut en detaljert reisedagbok (navn på destinasjon, nøyaktig adresse til destinasjon eller skjæringspunkt, reisemiddel, avgang- og ankomsttider, målsetting med reisen) for en dag i løpet av studien. I tillegg bad man om kontaktinformasjon for respondenten, og man spurte etter eventuelt eierskap til bil og om å oppgi steder man besøkte ofte. Poenget med å be informantene om å fylle ut en reisedagbok i tillegg til GPS-sporingen var å skulle kunne sammenligne disse to informasjonskildene, og slik generere et «truth dataset» som forteller i hvilken grad en GPS-basert studie kan forventes å gi en presis redegjørelse for individuelle reisevaner. Dette datasettet kan brukes andre veien også, i den forstand at det kan nyttes til å gjøre rede for i hvilken grad informantene til reisedagboken er i stand til å gjøre rede for reisene sine på en god måte i retrospekt.

Sneade (2013) gjennomfører en annen type datavalideringsundersøkelse. Her samles det inn GPS-data for et utvalg som fra National Travel Survey i Storbritannia. Husholdningsmedlemmer over 12 år bar med seg en GPS-enhet i 7 dager, og ble i tillegg bedt om å svare på en del spørsmål angående sine reisevaner. Denne studien spurte deretter respondentene om ulike forhold knyttet til det å bruke GPS-sporing i reisevaneundersøkelser. For eksempel holdninger til GPS-sporing, og opplevelsen av det å bli sporet med GPS. Studien konkluderer på bakgrunn av dette med at GPS ikke kan erstatte den tradisjonelle reisedagboken.

Roux m.fl. (2008) gjennomfører en studie av et begrenset utvalg fra den franske nasjonale reisevaneundersøkelsen.

2.4.1 GPS-baserte reisevaneundersøkelser

Som nevnt over har den mest utbredte bruken av GPS-sporing vært gjennomført i forbindelse med validering av datakvalitet for tradisjonelle reisevaneundersøkelser. Etter hvert er det gjennomført en del rene GPS-piloter, og noen få rene full-skala GPS-undersøkelser.

Stopher m.fl. (2013) gjennomfører en reisevaneundersøkelse i Cincinatti-regionen i USA, og baserer seg utelukkende på GPS-innsamlede data. Respondentene var tilfeldig utvalgte fra adresseregisteret, med overrepresentasjon av personer med adresse nærme regionens universiteter, og adresser med muligheten til å bruke kollektiv transport². Det opprinnelige utvalget var på 5564 husholdninger, og det endelige datasettet bestod til slutt av komplett reiseinformasjon for 2059 husholdninger.

Respondentene ble sporet i tre dager, og gjennomførte så en PR-survey, der de fikk tilgang til kart som viste reisene de hadde gjennomført, og fikk mulighet til å komme med korrigeringer der ukorrekt data var registrert. Respondentene ble underveis bedt om å gjøre rede for om de glemte enheten på turer eller hele dager i løpet av datainnsamlingsperioden. Algoritmene som er i bruk under denne studien, avdekker reisemiddel og reisemål i de aller fleste av tilfellene, man oppnådde et representativt utvalg i forhold til de fleste grupper, og man konkluderer med at bruk av GPS i reisevaneundersøkelser bør trappes opp.

Barbeau m.fl. (2008) tester mobiltelefonapplikasjonene TRAC-IT på 14 deltagere i 3 uker. Denne studien har som formål å spore deltagere for å så kunne påvirke individuelle reisevaner i retning kollektiv transport på et senere tidspunkt. Applikasjonen var basert på aktiv sporing, og brukerne ble bedt om å oppgi informasjon om ofte besøkte steder, reisemiddel, reisebelegg og destinasjon. En utfordring man opplevde var det at brukerne glemte å starte reisen i applikasjonen.

Stenneth (2011) gjennomfører en pilot som baseres utelukkende på GPS-data. Studien har som målsetting å utvikle algoritmer som avdekker reisemiddel på en god måte. Modellene som utvikles gjør i stor grad rede for reisene på en god måte, med opp til 93 % nøyaktighet. En av forutsetningene for denne studien er at man generelt i samfunnet deler mer og mer av sine digitale spor gjennom sosiale media. Gjennom at man kobler dette til reisemiddel kan man få informasjon om flere sider av reisene, for eksempel karbonavtrykk eller kalorier forbrent.

2.5 Litteratur

Barbeau, Sean J., Miguel A. Labrador, Nevine L. Georggi, Philip L. Winters og Rafale A. Perez (2008): TRAC-IT– A Software Architecture Supporting Simultaneous Travel Behavior Data Collection & Real-time Location-Based Services for GPS-Enabled Mobile Phones. Tampa, FL: Center for Urban Transportation Research, Department of Computer Science and Engineering, University of South Florida.

Bricka, Stacey (2009): "What Is Different About Non-Response in GPS-Aided Surveys?", i Bonnel, Patrick, Martin Lee-Gosselin, Johanna Zmud og Jean-Loup Madre (red.) (2009): Transport Survey Methods. Keeping Up with a Changing World. Bingley: Emerald Group Publishing Limited.

Chen, Cynthia, Hongmian Gong, Catherine Lawson og Evan Bialostozky (2010): Evaluating the feasibility of a passive travel survey collection in a complex urban environment: Lessons learned from the New York City case study, i Transportation Research Part A 44 (2010) 830–840.

Gould, Jane (2013): "Cell Phone Enabled Travel Surveys: The Medium Moves the Message", i Zmud, Johanna, Martin Lee-Gosselin, Marcela Munizaga og Juan Antonio Carrasco (red.) (2013): Transport Survey Methods: Best Practice for Decision Making. Bingley: Emerald Group Publishing Limited.

² Med andre ord et område som ikke kan sammenlignes med Oslo-området i forhold til kollektivtransportdekning.

- Nitsche, Philippe, Peter Widhalm, Simon Breuss og Peter Maurer (2012): A strategy on how to utilize smartphones for automatically reconstructing trips in travel surveys, i *Procedia – Social and Behavioral Sciences* 48 (2012), sider 1033-1046.
- Persad-Maharaj, Narin, Sean J. Barbeau, Miguel A. Labrador, Philip L. Winters, Rafael Pérez og Nevine Labib Georggi (2008): Real-time travel path prediction using GPS-enabled mobile phones. Paper presented at the 15th World Congress on Intelligent Transportation Systems, New York, New York, November 16-20, 2008.
- Roux, Sophie, Philippe Marchal og Jimmy Armoogum (2009): Acceptability of the use of new technologies by interviewees in surveys. INRETS (French National Institute for Transport and Safety Research), France.
- Smoreda, Zbigniew, Ana-Maria Oltenau-Raimond og Thomas Couronné (2013): Spatiotemporal Data from Mobile Phones for Personal Mobility Assessment, i Zmud, Johanna, Martin Lee-Gosselin, Marcela Munizaga og Juan Antonio Carrasco (2013): *Transport Survey Methods: Best Practice for Decision Making*. Bingley: Emerald Group Publishing Limited.
- Sneade, Abby (2013): Using Accelerometer Equipped GPS Devices in Place of Travel Diaries, i Zmud, Johanna, Martin Lee-Gosselin, Marcela Munizaga og Juan Antonio Carrasco (2013): *Transport Survey Methods: Best Practice for Decision Making*. Bingley: Emerald Group Publishing Limited.
- Stenneth, Leon O. (2011): *Transportation Mode Detection using Mobile Phones and GIS Information*. University of Illinois, Chicago Department of Computer Science.
- Stenneth, Leon O. og Philip S. Yu (2011): *Privacy-Aware Mobile Location-Based Systems*. University of Illinois, Chicago Department of Computer Science.
- Stopher, Peter R. (2009a): "The Travel Survey Toolkit: Where to From Here?" i Bonnel, Patrick, Martin Lee-Gosselin, Johanna Zmud og Jean-Loup Madre (red.) (2009): *Transport Survey Methods. Keeping Up with a Changing World*. Bingley: Emerald Group Publishing Limited.
- Stopher, Peter R. (2009b): "Collecting and Processing Data from Mobile Technologies", i Bonnel, Patrick, Martin Lee-Gosselin, Johanna Zmud og Jean-Loup Madre (red.) (2009): *Transport Survey Methods. Keeping Up with a Changing World*. Bingley: Emerald Group Publishing Limited.
- Stopher, Peter R., Christine Prasad, Laurie Wargelin og Jason Minser (2013): "Conducting a GPS-only household travel survey", i Zmud, Johanna, Martin Lee-Gosselin, Marcela Munizaga og Juan Antonio Carrasco (red.) (2013): *Transport Survey Methods: Best Practice for Decision Making*. Bingley: Emerald Group Publishing Limited.
- Wermuth, Manfred, Carsten Sommer og Marion Kretz (2003): "Impact of new technologies in travel surveys," i Stopher, Peter og Peter Jones (red.) (2003): *Transport survey, quality and innovation*. Oxford: Pergamon.
- Wolf, Jean, Michael Loechl, Miriam Thompson og Carlos Acre (2003): Trip rate analysis in GPS-enhanced personal travel surveys, i Stopher, Peter og Peter Jones (red.) (2003): *Transport survey, quality and innovation*. Oxford: Pergamon.

3 Teknikker for innsamling og analyse av data fra håndholdte enheter

3.1 Bakgrunn

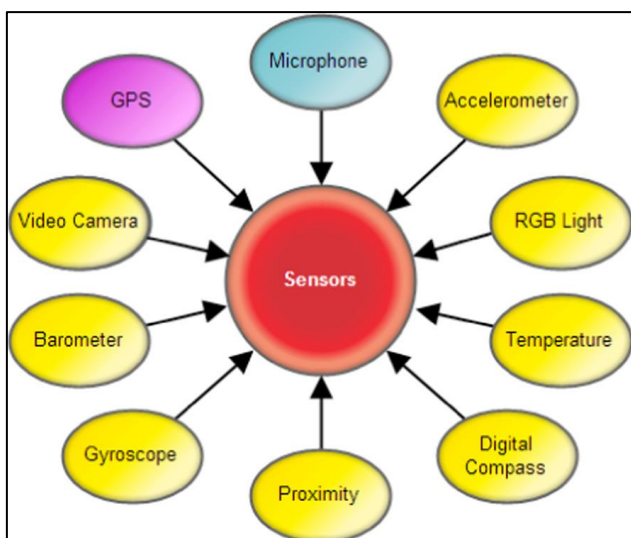
Dette kapittelet utfyller tidligere litteratursøk som er gjort rede for i kapittel 2 (Haugsbø 2013) vedrørende bruk av GPS i reisevaneundersøkelser. Tidligere litteratursøk har hatt fokus på bruk av data fra håndholdte enheter sammenlignet med tradisjonelle reisevaneundersøkelser. Dette litteratursøket har fokus på optimalisering av selve teknologien og teknikker for innsamling og barbearbeiding av data.

Før smarttelefoner ble vanlig, var det gjennomført flere forsøk med registrering med egne enheter. Flere av artiklene publisert før 2010 benyttet egne enheter med GPS for logging av lokalisering. Det har også vært et fokus på at bruk av GPS skal utfylle tradisjonelle reisevaneundersøkelser, blant annet med mer nøyaktige beskrivelse av reiser (Nitsche, Widhalm et al. 2012), (Xiao, Low et al. 2012), (Chen, Gong et al. 2010) og (Anderson, Abeywardana et al. 2009).

Etter 2010 har flere artikler fokusert på bruk av Smarttelefon for innsamling av data. Dette medfører noen utfordringer, men også flere fordeler. Blant utfordringene er noe dårligere GPS-mottak på grunn av plassering av smarttelefon i vesker, lommer mm. En annen utfordring er batterikapasitet og tømming av batteri i smarttelefonene.

Blant fordelene er at smarttelefonen bestandig er med og en trenger ingen egen enhet for logging. Registreringsenheten blir mindre synlig og påvirker hverdagen i mindre grad enn håndtering av en egen enhet. En annen fordel er introduksjonen av nye sensorer, muligheter for identifikasjon og kommunikasjonsmuligheter.

Det finnes muligheter for å benytte eget utstyr for å registrere og identifisere smarttelefoner med tid, sted og ID for passering over snitt. Notatet fokuserer på smarttelefonen og bruk av denne uten ekstra utstyr for å registrere eller identifisere brukerne. De største muligheter for kostnadsutt og nyttevurderinger er å unngå bruk av eget fysisk utstyr til registreringer. Utnyttelse av mulighetene smarttelefoner gir, kan erstatte vegkantutstyr (Herrera, Work et al. 2010).



Figur 3.1: Sensorer i en smarttelefon (Daponte, De Vito et al. 2013)

3.2 Innsamling av lokaliseringsdata

3.2.1 Utfordringer med smarttelefon

Smarttelefon er velegnet til å samle inn data. Samtidig skaper innsamling av data noen utfordringer. Batterikapasitet på smarttelefoner vil ofte bli påvirket dersom en samler lokaliseringsdata. I presentasjonen "How to track location in a battery-friendly way?" (Abdulazim 2013) blir diverse metoder diskutert:

- Bruk av nettverksbasert posisjon. (Dvs å bruke mobilen sin innebygde funksjon for å gjette posisjon basert på nettverk. Baserer seg ofte på hvilke telefonmaster eller trådløse nettverk en er koblet til.)
- Ta vare på "landemerker", for eksempel trådløse nettverk, for ofte besøkte posisjoner. (Dvs. å huske hvor et gitt trådløst nettverk er tilgjengelig - f.eks. hjemme -, og dermed vite hvor en befinner seg når dette nettverket er tilgjengelig.)

3.2.2 Mulighet for ikke-forstyrrende logging

En rekke tjenester har oppstått som gir brukeren nytte av å dele data, ofte omtalt som Crowd-Sourcing. Dette gir mulighet for å tilby tjenester basert på at en er villing til å dele egne data. Tidlig ble det utviklet tjenester som utnyttet mulighetene med smarttelefon og Crowd sourcing. Eksempler er (Cambridge Systematics 2012): Google Maps, Google Latitude, TomTom's MapShare, The Trapster application og Waze.

3.3 Bearbeiding av data

3.3.1 Analysemetoder

Flere metoder kan benyttes til å *gjenkjenne reisemønsteret* basert på maskinlæring som nevrale nettverk, Support Vector Machine og Random Forest. Hver av disse metodene er kompliserte og trenger større utledninger for å forklare. For enkelte vurderinger og eksempler på bruk kan en se f.eks. (Lin and Hsu 2014).

3.3.2 Reisemønster

3.3.2.1 Transportatferd og bruk av sosiale media

Bruk av sosiale media kan ha sammenheng med transportatferd. Sosioøkonomiske karakteristika og bruk av kommunikasjon påvirker både aktivitetsområde og hvor mye en reiser. Det understrekes at en bør være forsiktig med å konkludere, siden det kan stilles spørsmål med kausaliteten mellom enkelte av variablene. F.eks er antall turer avhengig av antall tekstbeskjeder eller omvendt (Ythier, Walker et al. 2013).

Det finnes også eksempler på prosjekt som prøver å utvikle *automatiske tekstlig beskrivelser av turer* (Frank, Mannor et al. 2013). Ved å bruke smarttelefonens sensorer og lokalisering, har en forsøkt å utvikle en modell for oversettelse av registreringer til en tekstlig beskrivelse av turen.

3.3.2.2 Klassifisering av reisemønster

En populær måte å kategorisere reisemønster på, er "Trajectory mining" (Lin and Hsu 2014). En viktig utfordring for "trajectory mining" er å finnes noen viktige punkt som kan beskrive reisen. "Trajectory mining" kan deles inn i to: "Trajectory clustering" som har til hensikt å finne felles *beskrivelse av bane eller rute* for en felles gruppe. Den andre - "trajectory patterns extraction" - brukes til å modellere felles adferdsparametere og *predikere fremtiden*.

Det er ofte en utfordring å se hvorfor forskjellige steder gir forskjellige reisemønstre. *Logging av ruter* kan gi en pekepinne på hvorfor disse forskjellene opptrer (Calabrese, Diao et al. 2013).

3.3.2.3 Trafikkmengde

Belegget på kollektivtrafikk kan ofte være vanskelig å registrere. I Paris (Aguilera, Allio et al. 2014) har en gjort forsøk med telefonnettverket og kapasiteten i nettverket til å *anslå antall passasjerer* som finnes på t-banen. Metoden forutsetter ikke gjenkjenning og telling av telefoner, men baserer seg på Quality of Service (QoS), i form av signalstyrke. Passasjerene er derfor helt anonyme.

3.3.3 Utledning av aktiviteter

3.3.3.1 Basert på data om lokalisering

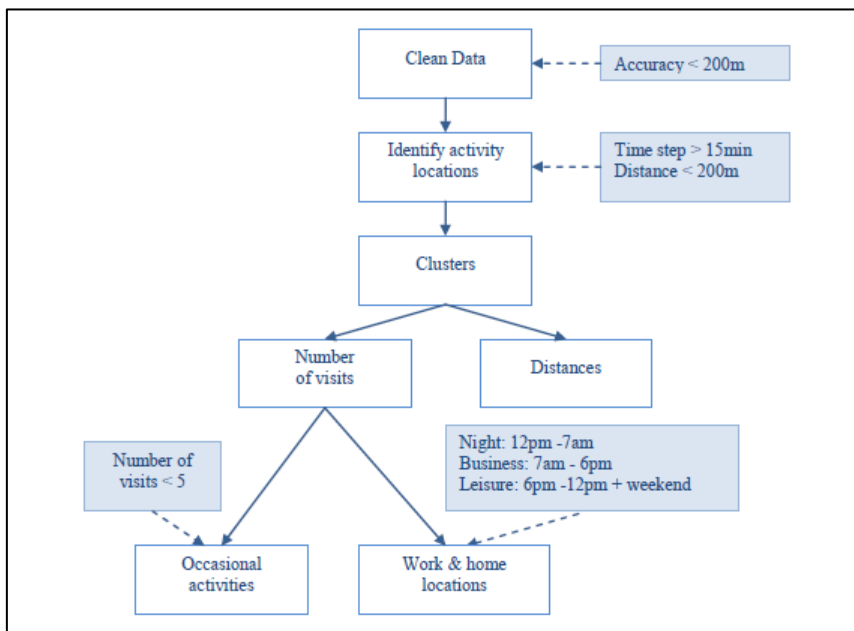
Lokaliseringsdata samlet inn fra mobile enheter kan foredles med andre datakilder for å avlede den reisendes aktiviteter. Ved bruk av *eksisterende kart og map-matching* (Wanga 2012) og (Bierlaire, Chen et al. 2013) kan en avlede gjeldende aktivitet. Eksempler på andre datakilder for å avlede den reisendes aktiviteter er "Land use information". ("How to infer traveller's activity with minimum burden?"):

Eksempler på tjenester som kombinerer lokasjonsdata fra brukeren med informasjon om stedet er Foursquare, Google Places, Yelp, Facebook og YellowPages.

En tjeneste som Foursquare har over 20 millioner brukere. De benytter en rekke datakilder med over 60 000 møtesteder. De mest benyttede møtestedene har 13 816 sjekk av lokalisering (Abdulazim 2013).

3.3.3.2 Basert på selvlæring

Type reisedestinasjon kan også avledes fra reisemønstre. Figuren under viser et eksempel på bruk av data for å avlede formålet med reisen.



Figur 3.2: Selvlæring av aktiviteter fra (Ythier, Walker et al. 2013).

Mønster-gjenkjenning benyttes også av (Lin and Hsu 2014) for å detektere *viktige lokaliseringer* for aktiviteter. Steder som blir hyppig besøkt av individer eller grupper som oppholder seg der over lenger tid, er ofte viktige for aktiviteter eller i analyser av transport. En kan avlede arbeidsplass for personer, og en kan koble data mot geografisk informasjon om parker osv. for å vise aktivitet. For transport kan en

også benytte metoder for mønstergjenkjenning til å analysere viktige parkeringsplasser basert på gjentakende mønster. Også *tap av GPS-signal* benyttes for å sannsynliggjøre viktige steder, siden dette ofte skyldes at en går inn i bygninger.

3.3.3.3 Prediksjon av aktivitet

Ved å kjenne igjen nåværende aktivitet og lokalisering i forhold til aktiviteter, kan en også gå videre til å analysere *sannsynlig fremtidig aktivitet*. I (Lin and Hsu 2014) og (Alvarez-Garcia, Ortega et al. 2010) er forskjellige metoder diskutert, blant annet er bruk av en Markov-modell beskrevet. Metoden tar utgangspunkt i bortfall av GPS-data, og baserer seg på sammenfallende grupper av registreringer på forskjellige områdenivå. Områdenivåene kan deretter detaljeres på mer detaljerte nivå og bevegelser mellom disse detaljerte nivåene kan analyseres videre. Siden hovedelementet i analysene er bortfall av GPS-data, vil hovedelementet være bygninger. Flere andre metoder er også diskutert i (Lin and Hsu 2014).

Prediksjon av ruter er også diskutert i (Chen, Lv et al. 2010). Metoden de foreslår, kombinerer prediksjon av rutevalg og destinasjon. Metoden bygger på at vegvalg ofte er repetitive. De utviklet en egen app som baserte seg på brukerens historiske data, og forutså fremtidig destinasjon og rutevalg.

3.3.4 Deteksjon av grupper

Smarttelefon kan benyttes til å detektere grupper av mennesker. I (Boston, Mardenfeld et al. 2014) diskuteres hvordan en kan benytte Bluetooth til å detektere tilstedeværelse av en gruppe mennesker. Ved bruk av Bluetooth kan en detektere *møtefrekvens og lengde på møter i grupper*. I tillegg til å gi mulighet for å *kjenne igjen grupper* (på arbeid, fritid osv) kan en utnytte dette selv om en ikke kjenner lokaliseringen. En kan på denne måten kjenne igjen et jobberangement arrangert på annet sted enn selve arbeidsplassen. En kan tenke seg at en kan gjennomføre dette med kobling på en sentral server eller desentralisert slik at hver enkelt enhet/mobil kan oppdage hvem som er i nærheten og i gruppen.

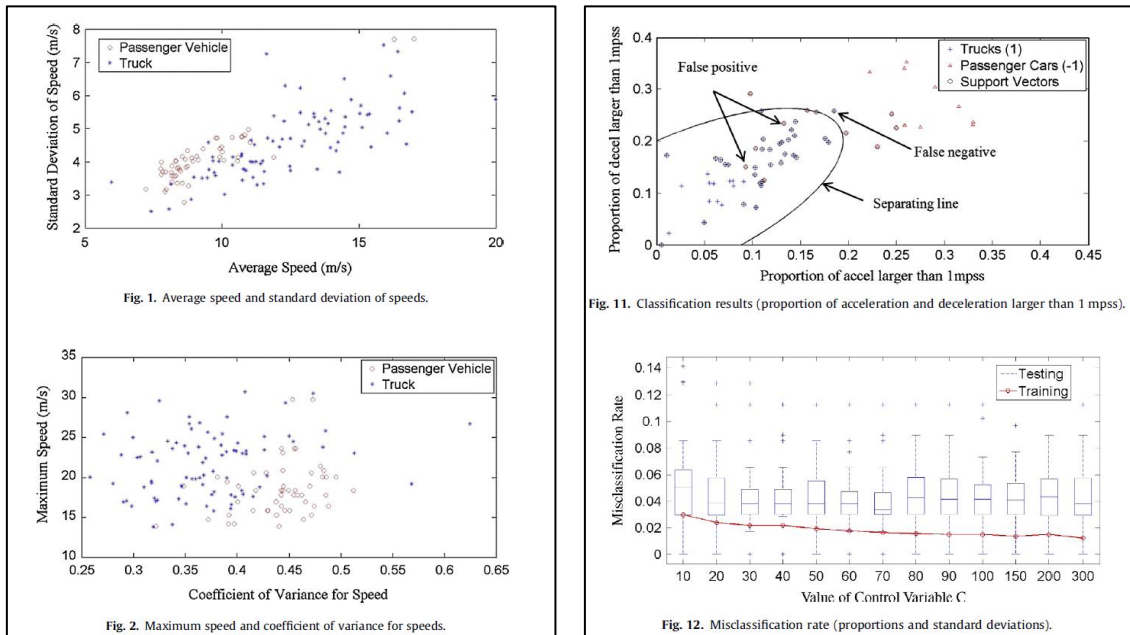
3.4 Utledning av transportmodi

I tillegg til reizens formål, kan mobiltelefoners mange sensorer også benyttes til å utlede transportmodi.

3.4.1 GPS-data

Flere kilder diskuterer hvordan en kan detektere transportmodi basert på GPS-data. I (Lin and Hsu 2014) diskuteres flere av disse metodene. Preprosesseringen er ofte ganske lik i forskjellige modeller og består i å *identifisere punkt for bytte av transportmodi* (deteksjon av felles punkt for bytte av transportmodi) og å *kjenne igjen transportmodi* basert på turens egenskaper. En viktig utfordring er å finne steder der det skjer *bytte mellom gange og andre transportformer*. Det er dog flere metoder som benyttes for å videre-behandle dataene for å kjenne igjen transportmodi. Fart, fartsendringer, snittfart, retningsforandring og antall stopp er indikatorer som benyttes i disse metodene.

GPS-data er ikke bare brukt til å klassifisere transportmodi som bil, buss eller lignende, men også til kjøretøyklassifisering. Spesielt data om akselerasjon og retardasjon kan benyttes for å studere kjøretøytype (Sun and Ban 2013). Ved å benytte GPS-data mente de å kunne skille mellom tunge og lette kjøretøy med en nøyaktighet på bedre enn 5 %.

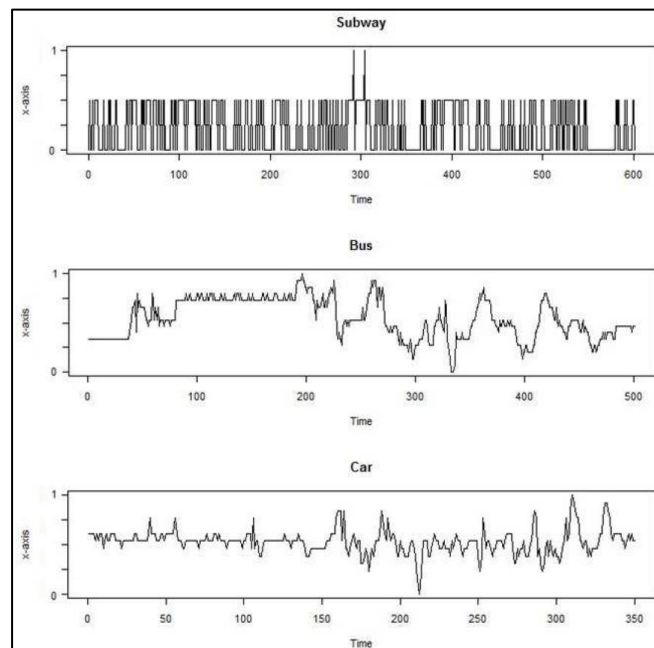


Figur 3.3: Eksempel på kjøretøyklassifisering fra (Sun and Ban 2013).

Enkelte transportmodi har det vært vanskelig å gjennomføre registreringer for, som f.eks. sykkel og gående. Bruk av smarttelefoner gjør det nå mulig å gjennomføre slike registreringer til f.eks. å finne sykkelruter (Hudson, Duthie et al. 2012).

3.4.2 Bevegelsessensorer

Etter at en startet å benytte smarttelefoner til registreringer, har en fått tilgang til en rekke sensorer som f.eks. akselerometer, gyroskop, og retning. Figuren under viser eksempel på registrering langs lengdeaksen i forhold til transportmodi.



Figur 3.4: Transportmodi basert på gyroskop fra (Abdulazim 2013)

I et eksperiment med å benytte gyroskop som sensor og maskinlæring (Abdulazim 2013), oppnådde en 98 % treff på transportmodi. Totalt så en på syv transportmodi; gange, løping, sykkel, bil, buss, undergrunnsbane.

3.4.3 GPS med supplerende data

Kombinerer en data fra forskjellige sensorer, kan en oppnå viktige fordeler. GPS-sensorer på smarttelefoner kan være av dårlig kvalitet. Smarttelefoner er spesielt utsatt for å skjermes i vesker, av klær eller av andre hindre. Flere metoder for bruk av GPS-data forutsetter *kontinuerlig logging* av data for å garantere høy kvalitet på posisjons- og fartsdata. I tillegg medfører bruk av kun GPS-data, begrensninger i muligheten til å skille mellom å stå stille, gange, og motorisert. For å omgå dette problemet, foreslår en i (Nitsche, Widhalm et al. 2014) og (Feng and Timmermans 2013) å utnytte posisjonsdata fra mobilnettet og bruk av akselerometerdata fra smarttelefonen. Dette medfører blant annet at en kan differensiere mer nøyaktig mellom flere transportmodi. På denne måten kan en også registrere turer med svake eller uten GPS-signal som en tidligere måtte filtrere ut på grunn av datamangel. I det ene prosjektet (Nitsche, Widhalm et al. 2014) ble en egen app (for Android) utviklet for å logge sensorer. Totalt ble det studert åtte transportmodi, inkludert undergrunnsbane.

3.5 Konklusjon

Bruk av smarttelefon gir enormt mange muligheter for registreringer og analyse av data. I dette notatet er det diskutert hvordan en kan finne sammenhenger mellom aktivitet på sosiale media og transportomfang, hvordan en kan benytte sensorer i telefonen til å kjenne igjen andre personer og grupper til å registrere mer tradisjonelle data om reisemønster og transportmodi. Går en videre med analysene, kan en også predikere reiser og destinasjoner. De nye metodene for å registrere og predikere reiser, gir nye muligheter innenfor kort- og langsiktig transportplanlegging.

3.6 Litteratur

- Abdulazim, T. (2013). Connected Traveller: Advanced Activity, Vehicle Data and Information Services Using Smartphones.
- Aguilera, V., S. Allio, et al. (2014). "Using cell phone data to measure quality of service and passenger flows of Paris transit system." *Transportation Research Part C: Emerging Technologies* 43, Part 2(0): 198-211.
- Alvarez-Garcia, J. A., J. A. Ortega, et al. (2010). "Trip destination prediction based on past GPS log using a Hidden Markov Model." *Expert Systems with Applications* 37(12): 8166-8171.
- Anderson, T., V. Abeywardana, et al. (2009). National travel survey GPS feasibility study: Final report. . U. Department of Transport, Department of Transport, UK.
- Bierlaire, M., J. Chen, et al. (2013). "A probabilistic map matching method for smartphone GPS data." *Transportation Research Part C: Emerging Technologies* 26(0): 78-98.
- Boston, D., S. Mardenfeld, et al. (2014). "Leveraging Bluetooth co-location traces in group discovery algorithms." *Pervasive and Mobile Computing* 11(0): 88-105.
- Calabrese, F., M. Diao, et al. (2013). "Understanding individual mobility patterns from urban sensing data: A mobile phone trace example." *Transportation Research Part C: Emerging Technologies* 26(0): 301-313.
- Cambridge Systematics, I. (2012). Travel Time Data Collection. D. I. Florida Department of Transportation.

- Chen, C., H. Gong, et al. (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 Part A: Policy and Practice* 44(10): 830-840.
- Chen, L., M. Lv, et al. (2010). "A system for destination and future route prediction based on trajectory mining." *Pervasive and Mobile Computing* 6(6): 657-676.
- Daponte, P., L. De Vito, et al. (2013). "State of the art and future developments of measurement applications on smartphones." *Measurement* 46(9): 3291-3307.
- Feng, T. and H. J. P. Timmermans (2013). "Transportation mode recognition using GPS and accelerometer data." *Transportation Research Part C: Emerging Technologies* 37(0): 118-130.
- Frank, J., S. Mannor, et al. (2013). "Generating storylines from sensor data." *Pervasive and Mobile Computing* 9(6): 838-847.
- Haugsbø, M. S. (2013). *Bruk av GPS i reisevaneundersøkelser*.
- Herrera, J. C., D. B. Work, et al. (2010). "Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment." *Transportation Research Part C: Emerging Technologies* 18(4): 568-583.
- Hudson, J. G., J. C. Duthie, et al. (2012). *Using Smartphones to Collect Bicycle Travel Data in Texas*.
- Lin, M. and W.-J. Hsu (2014). "Mining GPS data for mobility patterns: A survey." *Pervasive and Mobile Computing* 12(0): 1-16.
- Nitsche, P., P. Widhalm, et al. (2014). "Supporting large-scale travel surveys with smartphones – A practical approach." *Transportation Research Part C: Emerging Technologies* 43, Part 2(0): 212-221.
- Nitsche, P., P. Widhalm, et al. (2012). "A Strategy on How to Utilize Smartphones for Automatically Reconstructing Trips in Travel Surveys." *Procedia - Social and Behavioral Sciences* 48(0): 1033-1046.
- Sun, Z. and X. Ban (2013). "Vehicle classification using GPS data." *Transportation Research Part C: Emerging Technologies* 37(0): 102-117.
- Wanga, J. (2012). "Design of Mobile Phone Networks and Map-matching Navigation System." *Procedia Engineering* 29(0): 2045-2049.
- Xiao, Y., D. Low, et al. (2012). *Transportation activity analysis using smartphones. Consumer Communications and Networking Conference (CCNC), 2012 IEEE, IEEE*.
- Ythier, J., J. L. Walker, et al. (2013). *The influence of social contacts and communication use on travel behavior: A smartphone-based study. Swiss Transport Research Conference. Monte Verità / Ascona*.

4 Oppsummering

Denne delen oppsummerer hovedkonklusjonene fra de to bidragene Haugsbø (2013) og Engen (2014). Oppsummert er det to aspekter ved bruk av GPS til innsamling av reisevanedata som diskuteres.

- Teknologiske utfordringer og muligheter
- Kjennetegn ved utvalget i GPS-basert undersøkelser

4.1 Ulike teknologier til anvendelse – utfordringer og muligheter

Allerede før smarttelefonen ble allemannseie ble det gjennomført reisevaneundersøkelser der data ble samlet inn gjennom sporing av reiser ved hjelp av GPS. Studiene ble da gjennomført ved hjelp av egne enheter for registrering, det er imidlertid mange fordeler med å bruke individenes egne smarttelefoner i slike studier:

- Stor andel av befolkningen som har en smarttelefon
- Telefonen er bestandig med
- Registreringen blir mindre synlig for respondenten
- Rask teknologiutvikling på dette området

Det eksisterer i dag enorme muligheter forbundet med tilgjengelig teknologi til sporing av reiser basert på lokasjonsdata. Under er det angitt noen hovedpunkter:

Innsamling av lokasjonsdata:

- Bruk av nettverksbasert informasjon – hindrer tapping av batteriet
- Ta vare på landemerker – hindrer tapping av batteriet
- Crowd-Sourcing – muliggjør passiv logging

Bearbeiding av data:

- Support Vector Machine og Random Forest – maskinbasert gjenkjenning av reisemønster
- Bruk av stedsangivelse i sosiale media
- Modell for oversettelse av registreringer til en tekstlig beskrivelse av turen
- Trajectory mining – kategorisering av reisemønster
- Bruk av kunnskap om kapasiteten i telefonnettverket til å anslå passasjerer på et transportmiddel

Utledning av aktiviteter:

- Bruk av andre lokasjonsdataverktøy til å avlede aktiviteter – for eksempel Google Places og Facebook
- Utledning av reisedestinasjon fra reisemønster:
- Bruk av Markov-modell til å analysere sannsynlig fremtidig aktivitet
- Bruk av Bluetooth til å detektere grupper

Utledning av transportmodi:

- Preprosessering – gjennom identifisering av punkt for bytte av transportmodi og bytte mellom transportmodi – basert på turens egenskaper
- Videre-behandling av data til å utlede transportmodi– bruk av fart, fartsendringer, snittfart, retningsforandring og antall stopp
- Bruk av akselerasjon og retardasjon til å utlede kjøretøytype – skiller mellom tunge og lette kjøretøy
- Bruk av bevegelsessensor til å utlede transportmodi
- Kombinasjonsbruk av informasjon fra smarttelefonens akselerometer og posisjonsdata fra mobilnettet

Utfordringer relatert til bruk av smarttelefoner til sporing av reiser ved hjelp av mobiltelefon:

- Smarttelefonen ligger gjerne plassert i en veske eller en lomme og får svakere GPS-signal som følge av dette
- Batterikapasiteten kan være dårlig

4.2 Kjennetegn ved utvalget i en GPS-basert reisevaneundersøkelse

Bruk av GPS-verktøy i reisevaneundersøkelser representerer muligheter til å øke både de kvalitative og kvantitative sidene ved data som samles inn. Tradisjonelt har reisevaneundersøkelser vært basert på utfylling av en reisedagbok der respondentens hukommelse og engasjement i å gi sannferdige opplysninger har påvirket datas kvalitet i stor grad. Videre vet man at denne måten å kartlegge reiser på oppleves som belastende for respondentene, gjennom at de må gjøre en relativt stor egeninnsats. Det er derfor fare for at mange ikke deltar i slike undersøkelser på grunn av stor brukerbelastning, og videre at de som faktisk deltar ikke gir et helt riktig bilde av situasjonen. Bruk av GPS i reisevaneundersøkelser gir helt konkret følgende fordeler:

- Bedre kartlegging av ulike reiseelementer (gange, sykkel, parkering, bytte- og transittider)
- Og relatert til dette unngår man problemene forbundet med respondentenes manglende forståelse av reisekjeder
- Større detaljrikdom i datamaterialet
 - o Mer nøyaktige data
 - o Data fra en lengre tidsperiode

Data er med andre ord ikke befestet med feilkilder som følge av respondentenes glemsel eller uvilje til å bruke tid på å samle inn data, men det er samtidig viktig å være klar over eventuelle nye feilkilder i datamaterialet. Det største ankepunktet relatert til denne typen datainnsamling er relatert til det at man begrenser undersøkelsens univers til den delen av befolkningen som har en smarttelefon, og man får følgelig nye representativitetsproblemer. Denne problemstillingen vil imidlertid bli mindre med tiden etterhvert som større og større deler av befolkningen skaffer seg en smarttelefon.

Data fra slike undersøkelser vil være befestet med følgende feilkilder relatert til både brukeraksept og brukerbelastning. Problemer med brukeraksept vil være relatert til det faktum at det er mer inngripende i den personlige sfæren å spore sine reisevaner samtidig og gjennom bruk av teknologiske hjelpemidler, sammenlignet med å gjøre rede for de under et intervju i etterkant. På denne måten kan det tenkes at en større del av befolkningen vil være negative til å delta i en slik undersøkelse. Videre er det ikke alle grupper som har like gode forutsetninger for å ta i bruk teknologien som kreves for å delta i slike undersøkelser. Problemer relatert til brukerbelastning er for eksempel relatert til sameksistens med den daglige bruken av mobiltelefonen og batterikapasitet og merarbeid for deltagerne gjennom at man må skru på applikasjonen før hver reise.

Som en motvekt til det at respondentene kan oppleve sporingen som belastende diskuteres det i litteraturen ulike insentiver som skal gjøre det attraktivt å delta i en slik undersøkelse:

- Å betale for mobildata-kostnader for respondentene i perioden som undersøkelsen foregår
- Å levere samtidige trafikkoppdateringer til brukeren langs ulike predikerte ruter
- Å gi deltagerne informasjon relatert til reisene de har gjennomført, for eksempel karbonavtrykk eller antall kalorier forbrent
- Økonomiske insentiver

Disse to utfordringene ved data vil avhenge av i hvilken grad data blir samlet inn passivt eller aktivt. Ved passiv sporing vil det ikke kreves medvirkning fra respondentene i noen særlig grad etter at man har samtykket til deltagelse. På denne måten vil brukerbelastningen bli minimal, men på den andre siden vil

man oppleve problemer knyttet til brukeraksept, da dette er særlig inngripende i den personlige sfære. I det motsatte tilfellet, ved aktiv sporing vil problemet med brukeraksept bli mindre, da respondentene selv velger når de skal la seg spore, mens man her vil oppleve problemer relatert til brukerbelastning.

Vedlegg 1: Sammendrag utvalgte studier

Samtidig datainnsamling og overføring av sanntidsinformasjon til bruker (Barbeau m.fl. 2009)

TRAC-IT er en mobiltelefonapplikasjon som er utviklet og testet ved University of South Florida. TRAC-IT kan måle utgangspunkt, destinasjon, start- og sluttider for reisen, rute, fart og avstand. Gjennom samhandling med individet som blir observert kan man også inkludere informasjon om reisemiddel, reisebelegg og destinasjonen for reisen. For å gi individene et insentiv til å delta i slike undersøkelser tilbyr TRAC-IT personlige sanntidstrafikkmeldinger basert på individets tidligere reiseruter, gjennom en *Path Prediction* algoritme.

Målsettingen med å utvikle TRAC-IT var å gjøre det mulig å spore personlige bevegelser mellom ulike kommunikasjonsmiddel (bil, sykkel, buss osv.), og over lengre tidsperioder (f.eks. uker vs. dager). Fokus var på å utvikle en mobiltelefon-applikasjon som kunne brukes av flest mulig, samt å gi applikasjonen egenskaper som gjorde det attraktivt for folk flest å være med i slike undersøkelser. Sporing ved hjelp av TRAC-IT er hovedsakelig basert på GPS-signaler, men basestasjonene i mobilnettet blir brukt i skygeområder.

For å spare batteri og kapasitet regulerer TRAC-IT mengden informasjon som blir samlet inn og sendt videre til databasen. En *location-aware state machine* algoritme regulerer med hvilken frekvens lokaliseringinformasjon blir hentet inn fra GPS maskinvaren. Dersom ugyldig informasjon blir rapportert flere ganger på rad, eller dersom brukeren står i ro over en lengre periode, vil applikasjonen gradvis senke frekvensen på rapporteringen og gå inn i dvalemodus. På samme måte fører gyldig lokaliseringinformasjon til at applikasjonen øker frekvensen på rapporteringen.

TRAC-IT bruker en *critical point* algoritme for å avgjøre hvilken informasjon som skal sendes fra applikasjonen på mobiltelefonen til databasen. Informasjon som blir samlet inn mens brukeren står stille, eller beveger seg langs en rett linje, blir regnet som unødvendig, mens informasjon som er nødvendig for å rekonstruere en brukers reiserute blir referert til som *kritiske punkt*. For å sikre at man ikke regner målefeil som retningsendringer blir disse bare regnet som kritiske punkt dersom de overstiger en viss fartsgrense. Første og siste punkt i alle segment blir regnet som kritiske.

Den samtidige overføringen av GPS-informasjon blir koblet sammen med historiske reisedata eller med informasjon som deltakeren har oppgitt på forhånd, og det er derfor stort sett mulig å gjøre rede for reisen på en komplett måte uten å vente på brukers korreksjoner.

14 brukere deltok i en TRAC-IT felttest. TRAC-IT ble satt i aktivt modus; data ble bare overført dersom brukeren selv startet en reise i applikasjonen, og var innstilt på å etterspørre GPS-data hvert fjerde sekund. Både *location-aware state machine*- og *critical point* algoritmene var avskrudd, dette for å få testet ut applikasjonen med mest mulig data. I forkant av undersøkelsen ble brukerne bedt om å opplyse om steder de besøkte ofte, for å minimere arbeidet til brukerne under selve undersøkelsen, da disse plassene blir lagt inn som hurtigvalg i applikasjonen.

Brukerne bar med seg mobiltelefonene i to-tre uker, og de ble bedt om å rapportere reisene sine etter hvert som de ble gjennomført. Det ble ikke gitt instruksjoner om hvordan mobiltelefonen skulle oppbevares på reisene, dette fordi det er praktisk for brukeren å plassere mobiltelefonen som han eller hun selv ønsker, og fordi på denne måten fikk forskerne innsikt i hvordan mobiltelefonene tar imot og avgir signal når de blir oppbevart i en veske eller lignende. Etter hver gjennomførte tur fikk brukerne en e-post med informasjon om turen de hadde gjennomført, brukerne som hadde tilgang til Outlook fikk turene sine oppgitt i kalenderfunksjonen. Brukerne fikk mulighet til å godta eller avvise den informasjonen som var blitt samlet inn om dem, og til å komme med korrigeringer og utfyllende kommentarer.

317 turer utført av de 14 deltakerne ble registrerte i løpet av de tre månedene som forsøket varte. Av de 317 turene som ble registrerte var det bare sju som ble rapporterte som feilaktige av brukerne i etterkant. Fem av disse viste seg å være feil fra brukerne sin side, mens de to siste kom av feil registrering av slutt punkt på reisen som følge av svake GPS-signaler. Det ble også rapportert om problemer med å få startet logging av turer fra brukerne, og som en følge av dette var det et ukjent antall turer som ikke ble rapporterte i det heile tatt. Disse dataproblemene kommer selvfølgelig i tillegg til det faktum at man ikke vet hvor mange turer som ble glemt, i og med at man brukte aktiv sporing

Turer som ble utført til fots var preget av å ha store hull i overføringene av datapunkt, noe som ble knyttet til at brukerne gjerne hadde mobiltelefonen plassert i en veske eller sekk. Kvaliteten på data samlet inn fra reiser med buss var dokumenterte noe bedre enn gåturer, mens bilturer var de turene som var best dokumenterte, noe som ble forklart med at i bilen har man gjerne mobiltelefonen lett tilgjengelig på dashboard eller i passasjerstet, og derfor med god tilgang til GPS-signaler. På turer som av ulike grunner ikke ble sporet ble det som regel likevel dokumentert både start- og destinasjonspunkt via GPS, dette fordi brukerne da hadde mobiltelefonen ut i fri luft for å starte eller stoppe sporingen.

Den største utfordringen som ble rapportert fra brukerne var det å huske å starte en tur. Det blir derfor fremhevet som viktig å satse på utvikling av den passive versjonen av TRAC-IT, slik at denne feilkilden kan reduseres. Dette fordi man på denne måten ikke lenger er avhengig av at brukeren husker å registrere turen, da han eller hun blir sporet hele tiden frem til at applikasjonen blir skrudd av. Ved bruk av passiv sporing kommer man tilbake til problemstillingene som omhandler mobiltelefonen sin funksjonalitet ved siden av applikasjonen, inkludert batteriutfordringene, og ikke minst personvern.

GPS og automatisk genererte reisevanedata (Nitsche m.fl. 2012)

Denne artikkelen tar for seg testing av en smarttelefon-applikasjon i Wien i Østerrike. 14 personer samlet inn 266 timer med reisedata over en periode på to måneder, med målsetting om å videreutvikle transportmodellene i området. Der GPS-signal ikke var tilgjengelig baserte man seg på det mobile base-nettet og på akselerasjonssensorer for å gjøre rede for reisen. Det blir lagt vekt på at spredningen av smarttelefoner betyr et enormt potensiale for reisevaneundersøkelser, men det blir samtidig understreket at man må redusere arbeidet for deltakerne i slike undersøkelser til et minimum. På bakgrunn av dette har man utviklet et verktøy som i stor grad er i stand til å identifisere reisemiddel uten brukermedvirkning.

For å gjennomføre felttesten av applikasjonen valgte forskerne ut tre ulike typer mobiltelefoner, med ulike tilpasninger til bruk, for å sikre data som var representative for de fleste mobiltelefoner. Alle mobiltelefonene var utstyrt med en akselerasjonssensor, og hadde *Android* operativsystemer. Dataloggings-applikasjonen som ble testet i denne undersøkelsen registrerte steds- og fartsinformasjon ved høy frekvens, og man skilte mellom transportmetodene: gange, sykkel, motorsykkel, bil, buss, elektrisk trikk, undergrunn, tog og ventende. Hvilket reisemiddel brukeren brukte ble registrert automatisk av applikasjonen gjennom GPS-signalet og delvis gjennom akselerasjonssensoren. Dersom feil reisemiddel ble registrert, kunne brukeren gå inn og rette dette manuelt på en enkel måte. Fjorten personer deltok i undersøkelsen over en periode på to måneder. Brukerne måtte selv skru på applikasjonen og notere/sjekk reisemiddel når reisen begynte.

For å lokalisere brukeren benytter applikasjonen både informasjon fra GPS-signal, og der dette ikke var tilgjengelig, basestasjonene i mobilnettverket. Gjennom akselerasjonssensorer var applikasjonen i stand til å identifisere fire ulike faser i brukerens bevegelser: stopp, akselerasjon, deselerasjon og bevegelse i jevn fart. Disse dataene ble brukt til å identifisere reisestrekninger og reisemiddel. Hvor nøyaktig målt det enkelte punktet var, ble også registrert av applikasjonen.

Det viste seg at reiser som ble foretatt med sykkel eller gående var målt med en nøyaktighet på henholdsvis 98 og 92 prosent. Ved å kombinere reisemidlene trikk, tog og undergrunn, fikk man en kategori kalt «skinner» med en nøyaktighet på 80 prosent. Enkelte karakteristikk ved trikk, bane og tog-nettet i Wien gjør dette fornuftig. 36 prosent av reisene med undergrunnsbanen var blitt registrerte som gående, noe som har blitt

knyttet til at brukeren gikk inne i toget. 76 % av reisene med bil, og 77 % av reisene med buss ble automatisk registrerte med riktig reisemiddel av applikasjonen. For å øke nøyaktigheten i forhold til det å registrere riktig reisemiddel automatisk, blir det fremhevet som viktig å videreutvikle bruken av akselerasjonssensorer i slike studier, da denne teknologien supplerer GPS-signalene på en meget god måte.

Passiv GPS-basert datainnsamling i et komplekst urbant miljø (Chen m.fl. 2010)

Denne studien tester muligheten for å bruke passive GPS-baserte innsamlingsmetoder i reisemiddelundersøkelser i komplekse urbane miljøer. Dette gjøres gjennom GIS-algoritmer som automatisk kan avdekke transportmiddel og reisemål. Målsettingen er at deltakere i slike studier ikke skal måtte oppgi noe annet enn sosiodemografiske opplysninger. Slik reduseres belastningen på brukerne dramatisk, og som en følge av dette kan den enkelte spores over en lengre periode, og man kan få et bedre datagrunnlag.

Å gjennomføre slike undersøkelser i New York er utfordrende hovedsakelig på bakgrunn av: 1) urbane dalfører, 2) et ekstremt tett befolket område, med et ekstremt varierte arealbruksmønster, og 3) komplekse transportnettverk. Man har store områder som er utenfor GPS-dekning, og reisene som foretas er preget av geografien: korte og hyppige reiser og variert reisemiddelbruk, som gjør det utfordrende å registrere alle reisens detaljer.

Man har i denne studien testet en rekke GIS-algoritmer, som automatisk avdekker reisemiddel og antyder målsetting med reisen. Prosessen er delt inn i fire ulike deler. Første skritt var en transport-survey hvor to datasett ble samlet inn. I det første datasettet inngikk 25 tilsatte ved the New York Metropolitan Transportation Council, som ble bedt om å bære med seg en GPS-enhet gjennom en arbeidsdag. Det andre datasettet bestod av 24 studenter ved the City University of New York, som ble bedt om å bære med seg senderen over en periode på fem virkedager.

Enheten logget hver persons posisjon hvert femte sekund, inkludert; dato, tid, lengde- og breddegrad, fart osv. Hver respondent skrudde på GPS-enheten hver morgen, og skrudde den av hver kveld når man ikke skulle ut igjen. Hver enkelt ble videre bedt om å fylle ut en detaljert reisedagbok (navn på destinasjon, nøyaktig adresse til destinasjon eller skjæringspunkt, reisemiddel, avgang- og ankomsttider, målsetting med reisen) for en dag i løpet av studien. I tillegg bad man om kontaktinformasjon for respondenten, og man spurte etter eventuelt eierskap til bil og om å oppgi steder man besøkte ofte. Poenget med å be informantene om å fylle ut en reisedagbok i tillegg til GPS-sporingen var å skulle kunne sammenligne disse to informasjonskildene, og slik generere et «truth dataset» som forteller i hvilken grad en GPS-basert studie kan forventes å gi en presis redegjørelse for individuelle reisevaner. Dette datasettet kan brukes andre veien også, i den forstand at det kan nyttes til å gjøre rede for i hvilken grad informantene til reisedagboken er i stand til å gjøre rede for reisene sine på en god måte i retrospekt.

Andre steg i denne studien var å utarbeide ulike kart-lag, som representerte ulike transportmiddel, f.eks. veinettet, bussruter inkludert stopp, undergrunnsystemet inkludert innganger/utganger og pendlertogruter og -stopp. De ulike transportmidlene er som sagt organisert i ulike lag, men mange plasser vil de ulike lagene krysse hverandre i kartet, uten nødvendigvis å være knyttet sammen. Undergrunnsbanens nettverk har gjerne kryssende eller delvis sammenfallende ruter med bussnettet som ligger som et lag oppå, og videre er det mulig at en bilreise foretas i mer eller mindre samme rutetrase som bussen tar. Reisen blir derfor foretatt på samme geografiske punkt, men det er likevel ikke nødvendigvis kontakt. F.eks. en respondent som blir knyttet til et punkt som er identifisert til en undergrunns-nedgang, vil være knyttet til undergrunns-nettet helt til personen blir knyttet til et punkt hvor det er mulig å komme seg ut igjen. Gjennom å kombinere informasjonen fra de ulike kartlagene kan man utarbeide en database som på en enkel og grei måte kan nyttes til å identifisere reisemiddel og reisemål.

Det tredje steget i undersøkelsen er myntet på å identifisere reisemiddel. Dette gjøres ved å først identifisere ulike segmenter av reisen, basert på informasjonen fra GPS-punktene som er samlet inn. Gjennom å sjekke om det er hull i GPS-overføringene (et hull blir regnet som 120 sekunder eller 250 meter), kan man skille mellom undergrunns-transport og transport på gateplan, da hull betyr tap av GPS signal, som taler for undergrunn.

Neste steg er å identifisere aktuelle punkt hvor det er sannsynlig at bruker endret reisemiddel. Dette blir gjort gjennom at man antar at enhver reisemiddelendring er forbundet med en kort gåtur, som i denne studien blir definert som over et minutt. Etter at segmentene i reisen er identifisert forsøker man å knytte hvert segment til et reisemiddel, gjennom en algoritme basert på de ulike reisemiddeleenes kjennetegn.

Det fjerde steget identifiserer målsettingen med reisen. Ut i fra opprinnelsessteder og reisemål kan en arealbruksalgoritme koble destinasjonen med de aktivitetene som man vet foregår på dette geografiske punktet. Der slike automatiske koblinger ikke er mulig kan man sannsynliggjøre at brukeren hadde en gitt målsetting med turen ved å basere seg på tidspunkt, brukerens historiske reisevaner og arealbruk i området rundt punktet.

Samtidig Path Prediction ved hjelp av GPS og mobiltelefoner (Persad-Maharaj m.fl. 2008)

Denne studien har som målsetting å minimere belastningen på respondentene ved GPS-baserte undersøkelser. Både for å minimere trafikkfarer (ved bilkjøring) forbundet med slike studier, men også for å gjøre det mer attraktivt å i det hele tatt delta, gjennom at belastningen ikke føles for stor. Gjennom å ta i bruk informasjon om individets historiske reisedata for å predikere fremtidig reisevei. En Path Maker algoritme kartlegger detaljene rundt en reise, og samler denne statistikken i en GIS-database. Informasjonen blir så benyttet i en Path Prediction algoritme, til å anta den gjeldende reiseruten, basert på historiske reisedata.

Det blir i denne studien lagt vekt på å tilby respondenten goder for å delta i undersøkelsen. Historiske reisedata vil bli brukt til å tilby respondenten sanntidsoppdaterte trafikkmeldinger som er relevante for den aktuelle reiseruten. Kombinert med oppretting av en personlig profil på respondentene, kan man tilby et personlig tilpasset «nyhetsbrev» med informasjon om tilbud i området som respondenten er i ferd med å bevege seg inn i.

Generating probabilistic path observation from GPS data for route choice modeling (Chen m.fl. 2008)

I denne studien ble 100 mobiltelefoner delt ut til respondenter, og tatt i bruk som respondentenes private telefon. Programvaren måler brukernes bevegelser automatisk, og det trengs derfor ikke brukermedvirkning. Data samles inn gjennom mange ulike funksjoner: kalender, lydopptak, GPS, wifi, akselerometer, lyssensor og kamera. I forhold til personvern inngår man en avtale med hver respondent i forhold til hvilke data man er interessert i å dele. I tillegg til informasjonen som ble samlet inn automatisk ble det gjennomført en nettbasert spørreundersøkelse. Brukerne blir konfrontert med reiser som er foretatt, gjennom at de vises i kart, og blir spurt om sine handlinger og valg rundt reisen. Modellen som blir brukt inneholder en algoritme som gjennom å ta i bruk GPS-sporet som er samlet inn predikerer ulike reiseruter som sporet kan høre til, i motsetning til andre slike studier som knytter GPS-sporet direkte til en kjent rute.

Brukeraksept av ny teknologi i spørreundersøkelser (Roux m.fl. 2009)

En gang i tiåret foretar det franske Samferdselsdepartementet og det Nasjonale statistikk instituttet en nasjonal husholdningsbasert reiseundersøkelse (FNIS). Målet med disse studiene er å beskrive korte og lengre reiser som er foretatt av franske husholdninger, og likeså bruk av privat og offentlig kommunikasjon. Studien har tre fokusområder: å beskrive reiser, eierskap til bil og bilbruk, tilgjengelighet til offentlige kommunikasjon. Seks ulike verktøy er i bruk for å samle inn data:

Første besøk i husholdningen:

- Et CAPI (computer-assisted personal interview) spørreskjema tas i bruk for å samle inn sosio-demografisk informasjon om respondentene i de utvalgte husholdningene.
- En reisedagbok over 7 dager tilegnes et av kjøretøyene i husholdningen.

Andre besøk i husholdningen:

- Man velger ut et familiemedlem over 16 år, ved hjelp av Kish-grid, som blir bedt om å beskrive de lengre reisene som er foretatt de siste tre månedene
- Samme person skal gjøre rede for sine reiser en ukedag og en helgedag før intervjuet

I tillegg:

- Et underutvalg, ca 1300 individer, vil gjøre rede for bruk av transportmiddel gjennom hele livsløpet sitt
- Et underutvalg, ca 750 frivillige, blir gitt en GPS-mottaker.

Undersøkelsen inneholder 19 000 respondenter, og innsamlingen av data skjer i seks ulike bølger, spredt over 12 måneder, for å nøytralisere sesongvariasjoner. Siden intervjuene uansett skal gjennomføres over to besøk, kan GPS-enhetene deles ut og samles inn i forbindelse med intervjuer. For å ikke utsette deltagerne for trafikkfare (ved bruk av bil) er det snakk om passiv sporing gjennom GPS-innretningen. Av personvern hensyn har brukerne muligheten til å selv velge om turene skal spores eller ikke. Ingen data blir overført samtidig, men logges, og blir overført til databasen etter at forskerteamet har hentet tilbake GPS-innretningen.

Datainnsamling over lenge tidsperioder er mulig da det er snakk om mindre byrde for respondenten. Eksempler på data som kan samles inn gjennom GPS, men som er vanskelig gjennom tradisjonelle reisedagbøker:

- Beskrivelser av veldig korte turer, glemmes ofte
- Rutevalg
- Presis informasjon om tidspunkt for påstigning og avstigning
- Beskrivelser av reiser som ble foretatt andre steder enn der man bor (for eksempel ferier)

Ulemper med GPS-basert datainnsamling er for eksempel batterikapasitet, signalmottak, eller at respondentene glemmer GPS-innretningen før reise, eller på andre måter er uforsiktige i bruken. Videre er det viktig å huske at GPS-baserte studier kan føre til et noe skjevt utvalg. FNTS inneholder et spørsmål om brukeraksept av bruk av GPS-sporing, med følgende svaralternativer:

- Yes, without condition
- Yes, if it is possible to turn it off
- Yes, but with other condition
- No

66 % av respondentene svarte at de ikke kunne tenke seg å bli utsatt for GPS-sporing. Ut av respondentene som sa seg villig til å delta i en slik undersøkelse avviste 19 % den foreslåtte GPS-enheten. Variabler som hadde innvirkning på om man kunne tenkt seg å delta i en GPS-basert undersøkelse:

- Inntekt, de med høyere inntekt er i større grad villige til å delta;
- Husholdningens eierskap til bil, et høyt antall biler korrelerer med deltakelse;
- Husholdninger med høyteknologisk utstyr (datamaskiner, mobiltelefoner) har høyere sannsynlighet for å delta
- Alder, yngre personer er mer entusiastiske til å være med i en GPS-basert studie;
- Kjønn, kvinner er mer motvillige til å delta
- Generell helse, personer med dårlig helse vil ikke delta
- Mobilitet (antall turer, korte og lange), personer med lav mobilitet er motvillige til å delta;
- Sosialt lag, personer som bor i tett befolkede områder har større sjanse for å delta i slike prosjekter.