

Trafikdage resume

Titelforslag:

Den lange vej til deling af anonymiserede Floating Car Data

Indlæggets baggrund og formål

I de sidste to årtier er en række forskning- og innovationsprojekter omhandlende GPS data, der beskriver kørselsadfærd i personbiler – de såkaldte ”Floating Car Data” (FCD), blevet gennemført. Mængden af FCD, der er blevet og fortsat bliver indsamlet er meget stor, men den fulde værdi af disse data bliver sjældent anvendt. De fleste datasæt er kun blevet analyseret indenfor et specifikt emneområde, og megen ekstra værdi kan antagelig opnås fra disse data, hvis de anvendes til et andet formål end det oprindelige. Enkeltstående eksempler på dette er bl.a. Intelligent Farttilpasningsprojekter, hvor deltagernes hastighedsvalg er blevet grundigt beskrevet, mens andre effekter såsom rejsetid kun er berørt af få studier. Ligeledes er der indsamlet meget store datamængder i forbindelse med videoregistreringer i kørende biler, som kun er analyseret i et begrænset omfang.

De manglende dataanalyser skyldes i mange tilfælde manglende ressourcer og/eller betydelig undervurdering af omkostningerne i forbindelse med dataanalyserne, mens det i andre tilfælde skyldes begrænsninger pga. datas indhold af personfølsomme oplysninger. Hvis data er anonymiserede, kan de til gengæld anvendes af tredje part til analyser. Udover dette, bemærkes det, at al dataindsamling finansieret af offentlige midler, principielt skal være tilgængelig for offentligheden jf. EU lov. I praksis sker datadeling dog sjældent, fordi få organisationer er villige til at dele deres data af forskellige årsager, herunder beskyttelse af personfølsomme oplysninger. På baggrund af ovenstående har det været ønsket, at FCD fra det danske forskningsprojekt, ITS Platform skal anonymiseres, så de kan anvendes af forskere og undervisere i ind- og udland. Data i ITS Platform blev indsamlet med en Onboard Unit (OBU), der var installeret i alle de deltagende biler. Et første forsøg på at anonymisere ITS Platform data i 2016-17 mislykkedes på grund af kompleksiteten af opgaven. Vores forskningsspørgsmål har derfor været: Kan vi anonymisere FCD, så centrale informationer om kørselsadfærd bibeholdes, uden at der bliver adgang til personfølsomme oplysninger?

Anvendte metoder, analyser og fremgangsmåde

For at tilstræbe tilstrækkelig anonymisering og mest mulig værdi af de anonymiserede data, er følgende arbejdsgang anvendt: Først er endepunkterne i alle ture blevet identificeret for at definere de enkelte ture. Dernæst er endepunktsklynger identificeret og en buffer med et tilfældigt centrum nær hver endepunktsklynge er anvendt til at fjerne FCD her indenfor. Desuden fjernes alle informationer, der binder ture sammen, dvs. at koble flere ture med samme person. Endvidere fjernes dato og tid for turene, der alene forbindes med tidspunkt på dagen samt hverdag/weekend (myldretid/nat etc.). I nærmere detaljer består anonymiseringen af denne fremgangsmåde:

- 1 FCD er opdelt i ture baseret på tid og OBU’ernes identitet. Små pauser i registreringen ignoreres som dataudfald (typisk 1-2 sekunder), mens pauser > 120 sekunder opfattes som en ny tur.
- 2 FCD er map matchet til Open Street Map.
- 3 Identifikation af endepunkterne (start og stop) for hver enkelt tur foretages.
- 4 Alle parkeringssituationer for hver individuel OBU (f.eks. start og stop) er grupperet, hvis afstanden er < 50 m.
- 5 Derefter beregner en algoritme den nødvendige radius af en cirkel, herunder en buffer (Buffer 1), som dækker mindst 50 adresser eller har en radius på 2.000 m. Adressetætheden

- i området definerer størrelsen af cirklen, men i områder med meget spredt bebyggelse er radiusstørrelsen sat til 2.000 m.
- 6 En tilfældig adresse inde i Buffer 1 vælges som centrum for en nye buffer (Buffer 2), hvor diameteren defineres af afstanden mellem periferien af Buffer 1 og centrum af Buffer 2.
 - 7 Alle FCD'er fra OBU'en fjernes fra Buffer 2, medmindre de passerer bufferen uden stop > 120 sekunder.
 - 8 På baggrund af turenes begyndelsestidspunkt fordeles turene som følger: Myldretid: (07-09 og 14-17), skulderperiode (09-14), aftenperiode (17-22) og "free flow" periode (22-07). Dette opdeles endvidere på hverdag og weekend.
 - 9 Endelig fjernes ID'et relateret til den unikke OBU. Det betyder, at den enkelte tur kan følges, men at forskellige ture ikke kan forbindes med hinanden umiddelbart.

Data

FCD blev samlet som en del af ITS Platform projektet fra 2012 til 2014. I alt 570 mio. positioner indgår som rådata i dette anonymiseringsarbejde. ITS Platform bestod af en OBU i de deltagende biler, en backend server og nogle applikationer. 425 primært privatejede biler (hvoraf 389 leverede data) var med. OBU'en var en mobil enhed, der var skræddersyet specielt til mobile ITS tjenester og FCD og kommunikation blev foretaget via GPRS. For at køre applikationerne og for at opnå kendskab til kørselsmønstre indsamlede OBU'erne FCD med 1 Hz med oplysninger om blandt andet: ID, position, tidspunkt, retning og hastighed.

Foreløbige resultater

Anonymiseringsproceduren har medført, at en del af FCD blev fjernet. De anonymiserede FCD består af ca. 409 mio. positioner, svarende til ca. 9,2 millioner kørte km, hvilket er en reduktion på 28% i forhold til rådata. Med afkortningen af turene i forbindelse med begyndelses- og slutpunkt reduceres turlængderne generelt, men da mange, især meget korte ture, fjernes helt, forøges den gennemsnitlige turlængde med 27%. Den gennemsnitlige turlængde blev forøget fra knapt 13 til 16 km.

De anonyme FCD vil komme til at ligge tilgængelige for alle på AAU's domæne i løbet af foråret 2019.

Forslag til emneplacering

Da artiklen vil arbejde med FCD kunne den emnemæssige placering være "Intelligente Transportløsninger". Det kunne dog også være en alternativ placering under emnet "Mobilitet og Adfærd". Skønt omdrejningspunktet i artiklen er FCD, giver efterfølgende analyser af disse data netop viden om dele af mobilitetsmønsteret blandt de, der har haft en OBU i deres bil. Det skal bemærkes, at en engelsksproget artikel med tilnærmelsesvist samme indhold er indsendt til ITS Europa Kongressen i Brainport i Holland, der afvikles i juni 2019.