Using GPS data as auxiliary data.

Anne Elevelt, Wim Bernasco, Peter Lugtig,
Stijn Ruiter & Vera Toepoel.

BigSurv18 conference. Barcelona. 27 October.

Where you at? Using GPS Locations in an Electronic Time Use Diary Study to Derive Functional Locations.

Anne Elevelt, Wim Bernasco, Peter Lugtig,
Stijn Ruiter & Vera Toepoel.

BigSurv18 conference. Barcelona. 27 October.

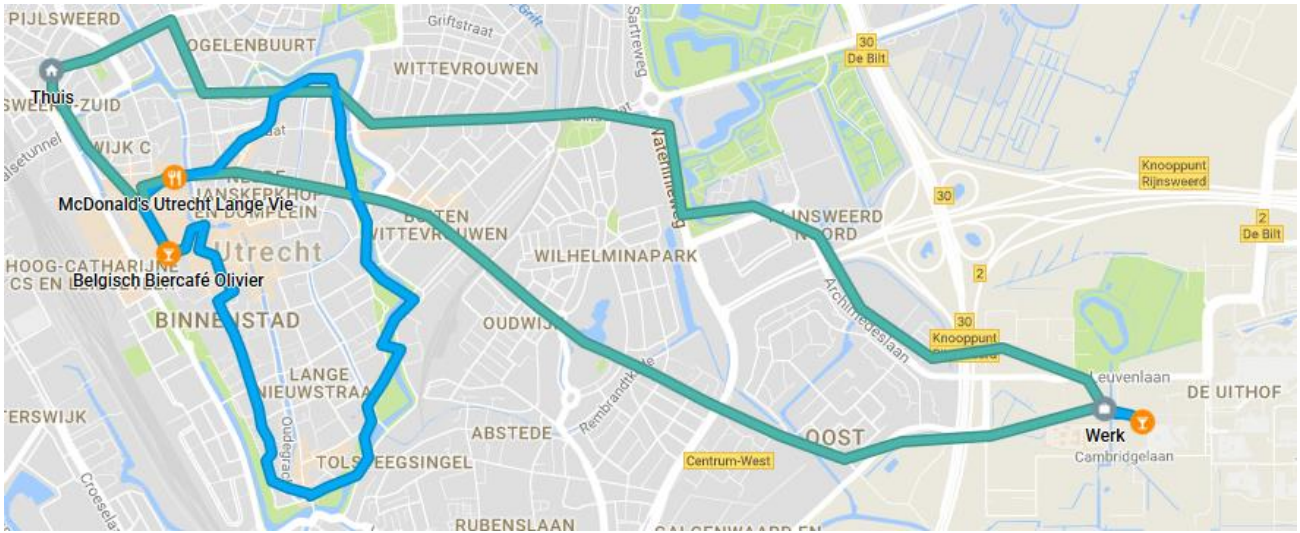
Diary studies

- Reliable and Accurate
- Different Domains
- But:
  - Burdensome

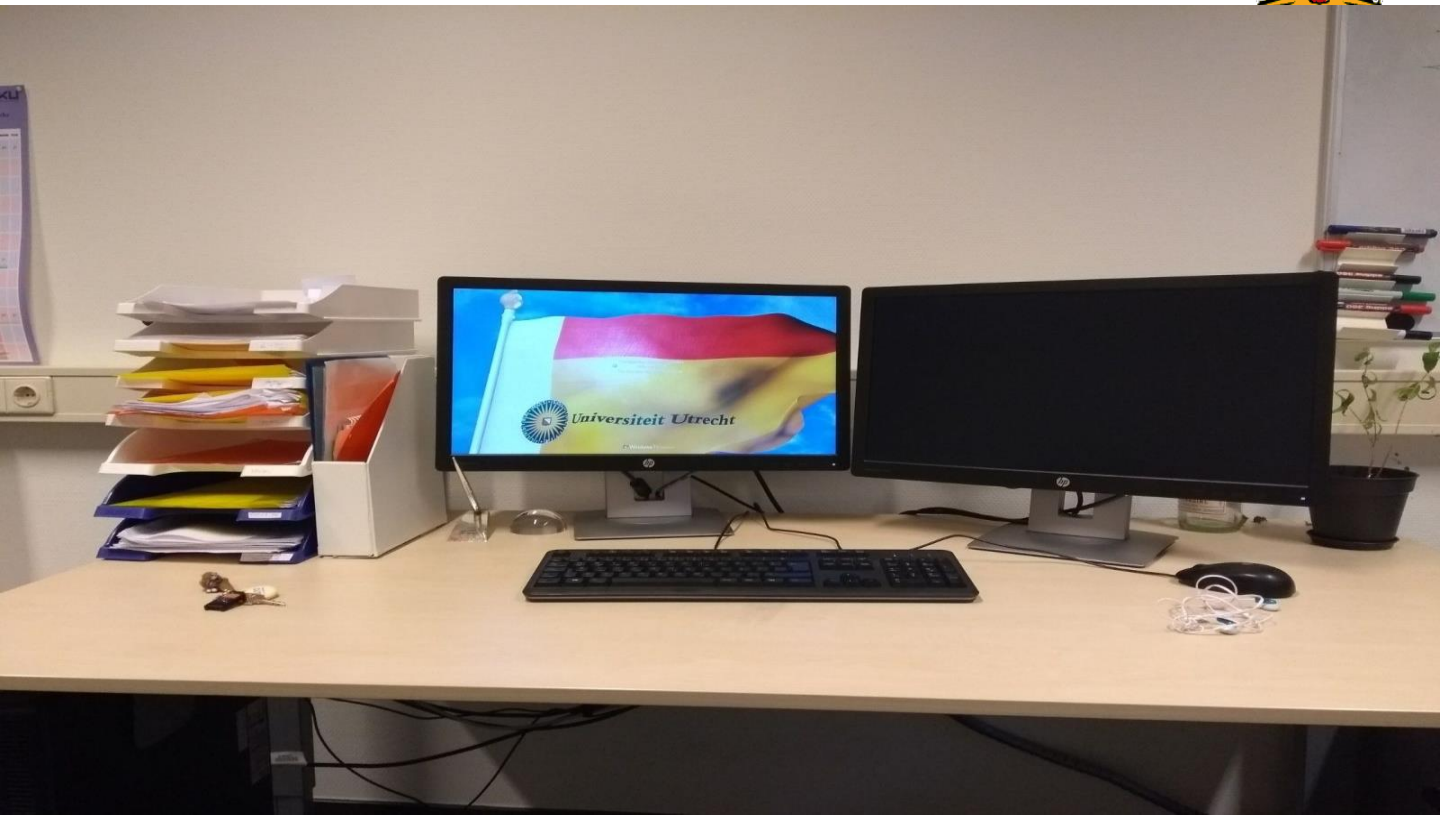
Time Use Survey App

For more information. see: Sonck & Fernee. 2013.

GPS tracks


PIJLSWEERD
Griftstraat
OGELENBUURT
WITTEVROUWEN
Sartrewg
30
De Bilt
Thuis
ZUID
WIJK C
Waterlinieweg
Knooppunt
Rijnsweerd
30
2
De Bilt
McDonald's Utrecht Lange Vie
Utrecht
WILHELMINAPARK
Belgisch Biercafé Olivier
BINNENSTAD
OUDWIJK
Archimedeslaan
30
Knooppunt
LANGE
NIEUWSTRAAT
Leuvenlaan
DE UITHOF
ABSTEDE
Rembrandtkade
OOST
Werk
Cambridgelaan
TOLSTEEGSINGEL
Centrum-West
Croeselaan
RUBENSLAAN

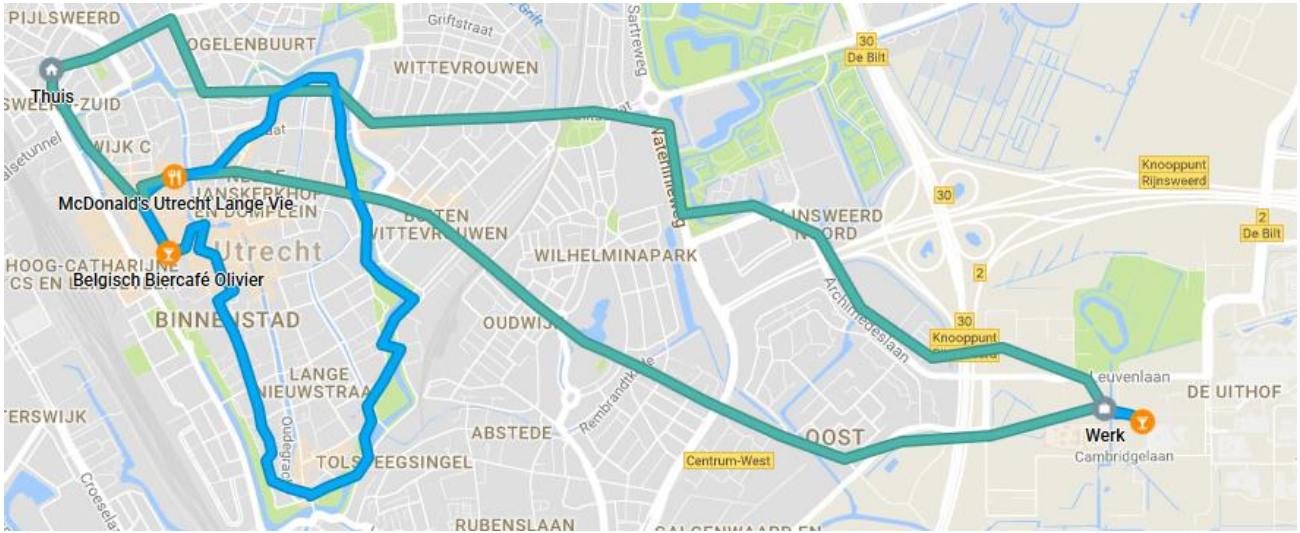

Without data, you're
just another person
with an opinion.


Universiteit Utrecht

Potential problems

- Battery life
- Data quality of GPS measurements
- Rule-based algorithms needed

GPS tracks.


Thuis
McDonald's Utrecht Lange Vie
Belgisch Biercafé Olivier
Werk
Cambridgelaan
Leuvenlaan
DE UITHOF
Knooppunt Rijnsweerd
Knooppunt
Centrum-West
WILHELMINAPARK
WITTEVROUWEN
ABSTEDE
RUBENSLAAN
Utrecht
BINNENSTAD
LANGE NIEUWSTRAAT
Griftstraat
Sartreweg
Archimedeslaan
Rembrandtlaan
30
2
De Bilt

GPS coordinates.

date	longitude	latitude
2018-02-23 17:44:14	5.124597	52.09303
2018-02-23 17:43:11	5.128350	52.09254
2018-02-23 17:42:09	5.130846	52.09203
2018-02-23 17:41:04	5.134006	52.09042
2018-02-23 17:40:03	5.136958	52.08891
2018-02-23 17:38:59	5.139979	52.08749
2018-02-23 17:37:56	5.143923	52.08639
2018-02-23 17:36:53	5.147190	52.08532

date	longitude	latitude
2018-02-23 08:21:42	5.170318	52.08520
2018-02-23 08:20:39	5.170312	52.08520
2018-02-23 08:19:30	5.170323	52.08520
2018-02-23 08:18:28	5.170312	52.08520
2018-02-23 08:17:26	5.170333	52.08523
2018-02-23 08:16:24	5.170320	52.08520
2018-02-23 08:15:22	5.170316	52.08520
2018-02-23 08:14:21	5.170304	52.08520
2018-02-23 08:13:17	5.170253	52.08519
2018-02-23 08:12:18	5.170216	52.08521
2018-02-23 08:10:57	5.170545	52.08503
2018-02-23 08:10:29	5.170417	52.08477
2018-02-23 08:09:16	5.169474	52.08611
2018-02-23 08:08:13	5.164519	52.08726
2018-02-23 08:07:12	5.160384	52.08687
2018-02-23 08:06:08	5.156182	52.08869
2018-02-23 08:05:07	5.153705	52.09132
2018-02-23 08:04:02	5.149904	52.09237
2018-02-23 08:03:00	5.145754	52.09208
2018-02-23 08:01:57	5.144957	52.09527
2018-02-23 08:00:54	5.140605	52.09573
2018-02-23 07:59:52	5.135603	52.09555
2018-02-23 07:58:49	5.130910	52.09543
2018-02-23 07:57:48	5.127493	52.09532
2018-02-23 07:56:45	5.125550	52.09649
2018-02-23 07:55:39	5.121906	52.09648
2018-02-23 07:54:36	5.117634	52.09642
2018-02-23 07:53:34	5.115756	52.09906
2018-02-23 07:52:02	5.111302	52.09772
2018-02-23 07:51:01	5.108787	52.09730
2018-02-23 07:49:59	5.108909	52.09726

Aims

- Propose a method for analyzing the GPS data
- Integrate the location and survey data
- Explain variability
- Understand how to use location data
  - Sufficient quality? -> Automatically record functional locations in the future

Functional Locations

- Home
- School / Work
- In Transit
- Everywhere
- Other

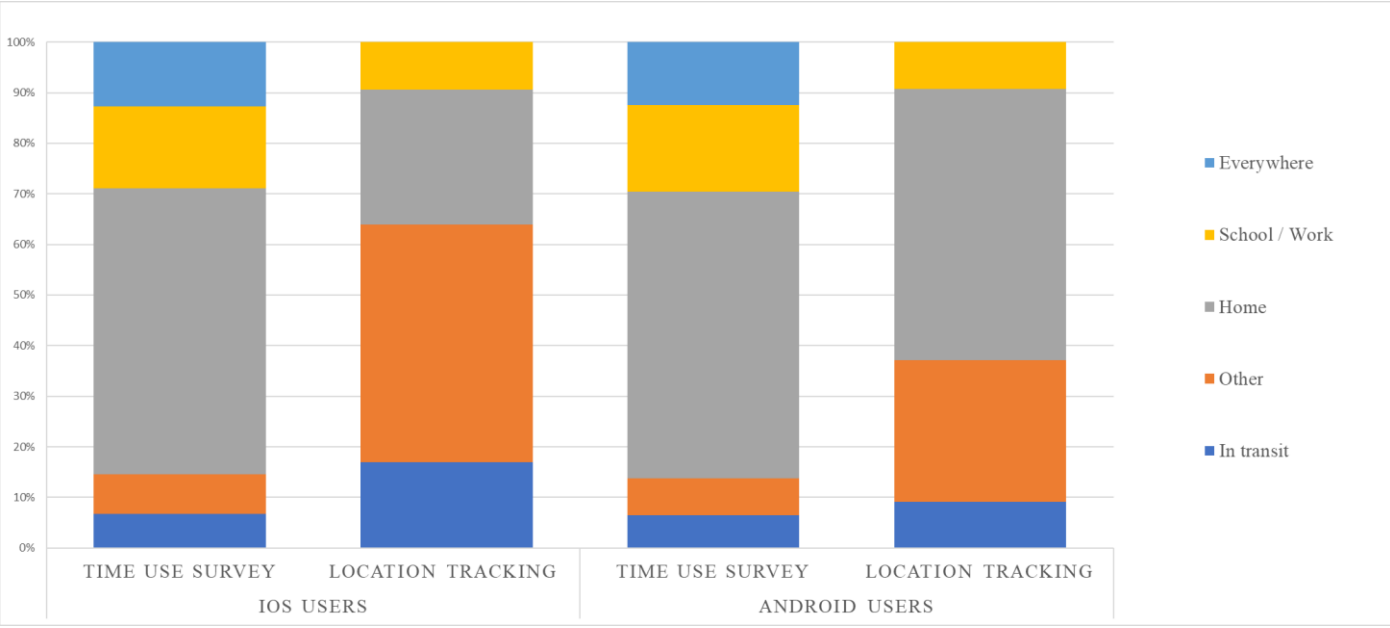
Do the two sources match?

- Overall distribution
- Timeslot level
- For what activities?
- Explore variability

Methodology

- Youth in Europe Study (YES!) panellists.
  - Around 19 years old
    - » Netherlands Institute for the Study of Crime and Law Enforcement [NSCR]
- Time Use Diary on Smartphones
  - GPS data
    - » Originally developed by CentERData and Netherlands Institude for Social Research [SCP].
    - » Sonck & Fernee. 2013.
- IOS (Iphone) versus Android users
  - Measurement Frequency

Overall Distribution



Timeslot level (Android)

		Based on Time Use Survey				
		In transit	Everywhere	Other	Home	Work
Based on GPS	In transit	**42.38**	10.23	15.73	3.76	10.17
	Other	33.08	33.37	**42.45**	22.04	35.46
	Home	20.05	49.40	34.71	**70.46**	22.39
	Work	4.49	7.00	7.11	3.75	**31.98**

Accuracy: 58.5%

Timeslot level (iOS Users)

		Based on Time Use Survey				
		In transit	Everywhere	Other	Home	Work
Based on GPS	In transit	**23.74**	19.89	22.93	12.79	24.07
	Other	53.91	52.44	**50.20**	46.03	41.23
	Home	11.20	18.26	16.60	**35.56**	13.40
	Work	11.16	9.41	10.27	5.61	**21.30**

Accuracy: 33.3%

Potential Problems *(and tested solutions)*

- Coding errors in location data
  - *Code over days*
- Coding errors in TUS data
- Data sources out of sync
  - *Allow small temporal inconsistencies*
- Not enough information in isolation
  - *Use information from TUS*

Time Use Categories

- No large differences within category
- iOS < Android
  - Sleeping: 41.3% vs 74.2%
  - Watching TV 19.6% vs 59.3%

Explore variability

Multilevel model, Android & iOS separately

MATCH <- Time of day + Weekend + times filled out

ICC: 0.121

Explore variability

Multilevel model, Complete sample

MATCH <- Android + Time of day + Weekend + times filled out

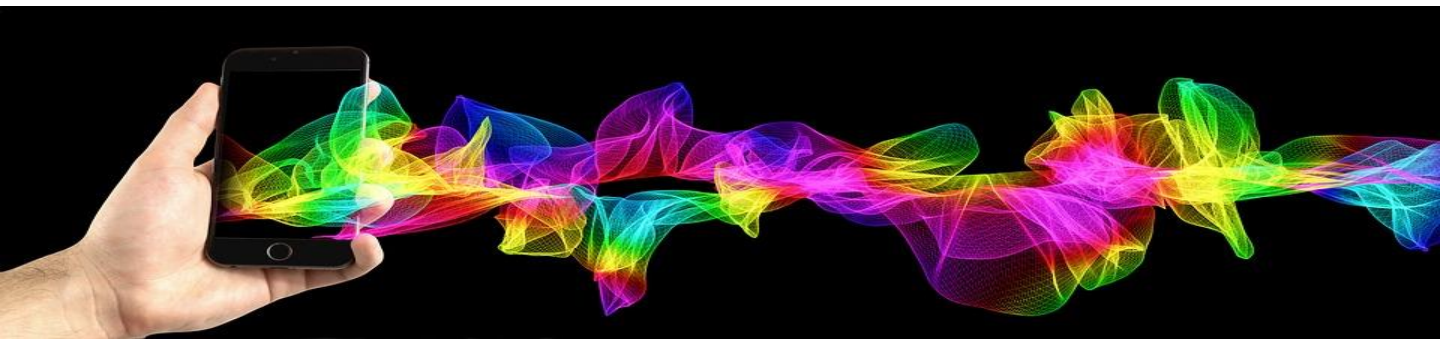
ICC: 0.142

Conclusions

- Many positioning errors.
- Differences in matching success are largely intrapersonal.
- Measurement frequency intertwined with phone type.

Next step

- Combine sensor and survey data!
- Experiment
- Positioning errors: Use better filter



Thank you!

Contact: a.elevelt@uu.nl

Twitter: @AnneElevelt