



MEILI, a small step towards automation of activity travel diary collection in Stockholm

Yusak O. Susilo
KTH Royal Institute of Technology
Transport Science Department
yusak.susilo@abe.kth.se



Acknowledgement

KTH Transport & Geoinformatics Depts.:

Adrian C. Prelipcean, Győző Gidófalvi

Sweco Society AB:

Andreas Allström, Ida Kristoffersson, Jenny Widell, Astrid Adelsköld, Edit Knutas

Linköping University:

Class Rydergren

TRV 2014/10422 (SPOT, Trialling and comparing smartphone based travel data collection with paper-and-pencil method)



Why do we do this?

- A good quality of data is paramount - to understand, predict, and anticipate
- Traditional survey methods are slow, expensive, and the quality of the data is poor (bias travel length & time, difficult to map the start- and end-point, forgotten/unreported trips, etc.)
- Especially for the Swedish national travel survey: dropping response rate (77% in 94/98, 42% in 11/14, 35% in 2015)
- Fast changing behaviour and shifting to activity based models will demand even more detailed data of higher quality



What do we need?

A way of summarizing **where**, **why** and **how** a user traveled during a the defined time frame by specifying:

- The (origin and) **destination** of a trip
- The (activity location and) trip's **purpose**
- The **means** of transportation (and their characteristics), i.e., trip legs and stages

Alternative method: passive & active mobile data collection, in particular: GPS + web and/or mobile-based data collection

Outsourced to/mining data from other providers?

White label?



MEILI: what is that?

MEILI

Verify Trip | Set Destination | Set Trip History | Proceed | Go Back

User: adi@kth.se | Logout

Select Language: English

Current Trip

Arrival Date: 9 September 2014
 Arrival Time: 10:13:56
 Departure Time: 09:17:21
 Destination:
 Mode: onfoot, subway, bus, onfoot

Trips left to process: 32

Previous Trips

Arrival Date: 7 September 2014
 Arrival Time: 13:16:17
 Departure Time: 13:11:45
 Destination: Oskarsrogatan 29 (home)
 Mode: onfoot

Arrival Date: 7 September 2014
 Arrival Time: 12:59:24
 Departure Time: 12:54:11
 Destination: Lidl (supermarket)
 Mode: onfoot

Arrival Date: 6 September 2014
 Arrival Time: 16:35:04
 Departure Time: 16:22:36
 Destination: Oskarsrogatan 29 (home)
 Mode: onfoot

of Odin,
 which data
 be easily
 es to collect
 d with
 readings
 ectioned data via
 and let users
 data into travel

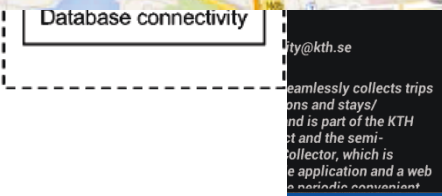
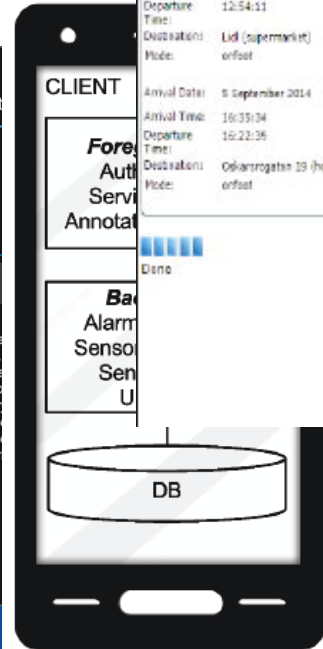
KTH Mobility

Username

Password

Register

The KTH Mobility Collection Pro application is voluntary user-controls for suspended a collected trip; and the analyzed in an anonymous purposes.



- updates
- store the annotations in a database



At the end of initial attempt

Almost perfect readings would kill the battery within 14-16 hours. MEILI adopts an *adaptive, equidistance, power-conscious* sampling strategy. It switches between two timer-triggered sampling loops: location sampling loop (when location readings are available) and accelerometer sampling loop (to initiate the location sampling loop when the accelerometer readings suggest movement). This make the unit able to last until 30 hours.

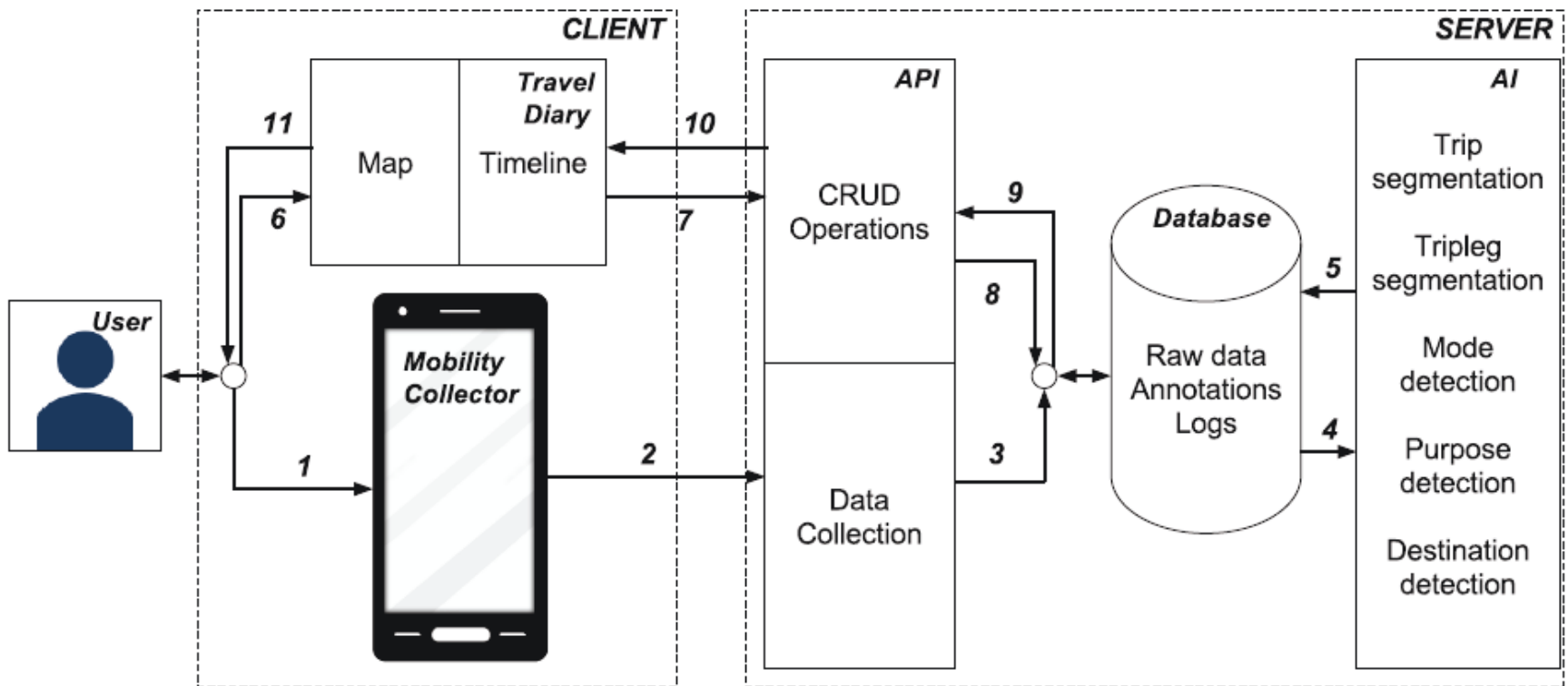
Android only implementation restricts the user pool in iOS predominant markets (such as Sweden)

Users wanted more freedom to interact with their data in the web interface
It was difficult to improve the system due to the lack of isolated functional components

It was difficult to extract trips and triplegs from a point-based model

Improvement based on initial attempt

Implemented Mobility Collector for iOS





The arrangement on the big trial:

1. Users install the MEILI Mobility Collector on their smartphones and start collecting data
2. After collecting data, the users log in to the MEILI Annotation website and annotate their trips by specifying:
 - The trip's destination
 - The trip's purpose
 - The triplets and their associated means of transportation
3. For comparison with PnP: on one day of the week, users have filled in a traditional travel survey



Resdagbok 4 november 2015

Parallellt med att du har en app installerad som samlar in data om hur du reser vill vi även att du svara på en traditionell resdagbok för dina resor under onsdagen den 4 november 2015. Detta behöver du bara göra för denna dag och syftet är att vi ska kunna jämföra de två metoderna. För att kunna jämföra de två metoderna kommer dina svar att sparas tillsammans med din e-postadress. Om du inte godkänner detta kontakta projektet via spot@sweco.se.

Webbased "Paper&pen" survey

1559 signed up
171 users used MEILI for at least one day
51 users used MEILI for at least one week
collected 2142 trips and 5961 triplegs
schema of 16 different travel modes
schema of 13 different purposes
POI set of 21953 entries in the database
transportation POI set of 6610 entries in the database

Hur gammal är du?

Välj alternativ: -- ▾

Var bor du?

Gata

Postnr (5 siffror utan mellanslag)



So what 1: What's the response rate?

What was the response rate in Stockholm RVU in September to October?

Trafikförvaltningen has answered that the response rate was 35%.

How many used the web-questionnaire instead of the paper questionnaire?

*19% of respondents used the web-questionnaire $\rightarrow 130\ 000 * 0.35 * 0.19 = 8645$. Out of these 1559 signed up for the SPOT-survey. However, the possibility to sign up for SPOT was introduced in the web-questionnaire in the beginning of the second week of the data collection period (3 weeks of data collection where $130000/5=26000$ users were contacted every third weekday, followed by 2 reminder weeks).*

Only 30 people signed up for SPOT because of the text in the paper questionnaire.



MEILI

On: 2 November 2015 - 11 November 2015 (**9 days**)

Annotated trips: 2142

Users: 171, annotated for at least 1 day

Paper and Pen

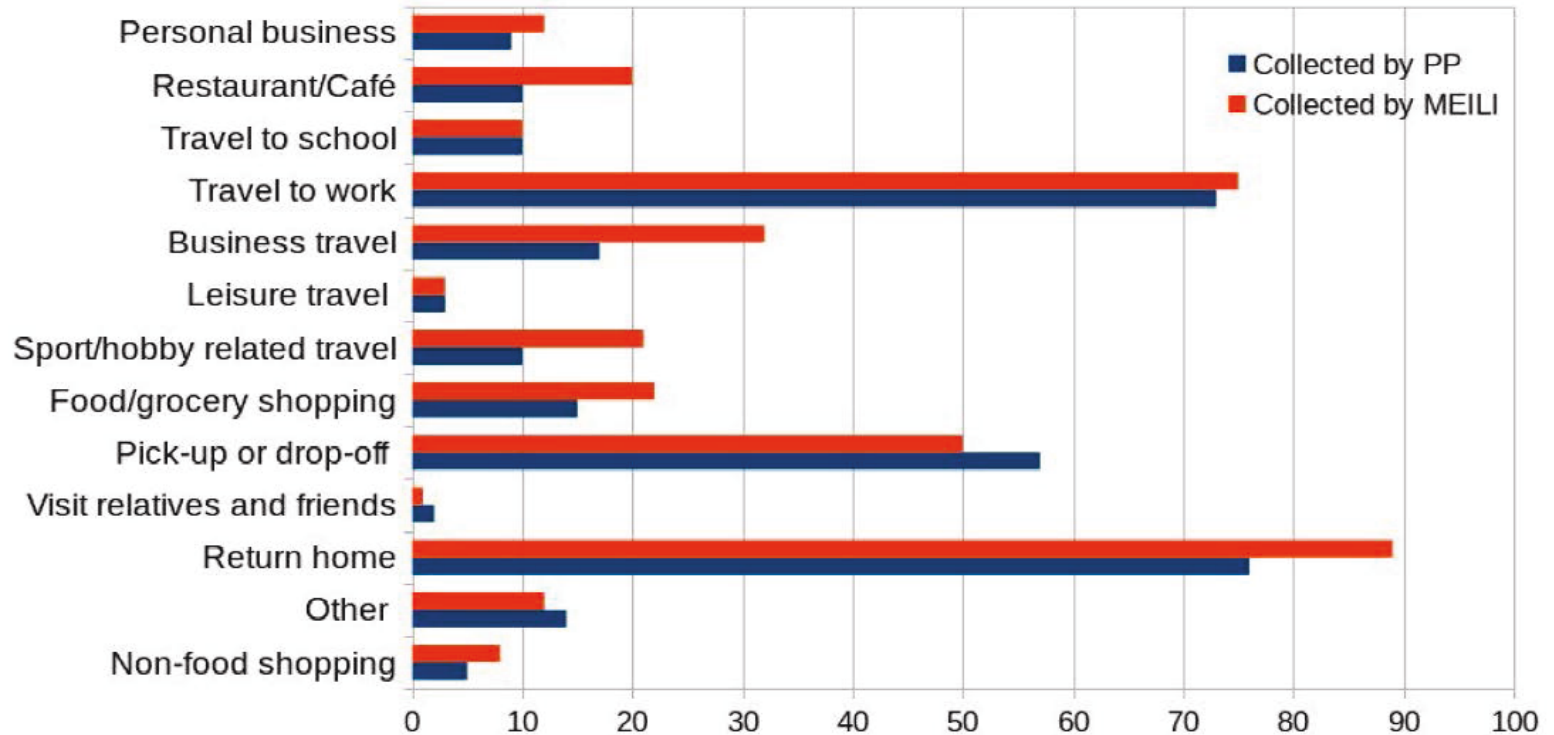
On: 4 November 2015 (**1 day**)

Users: 415 (83 users of them also reported their trips via MEILI)

Total number of trips: 1375 (301 from these same 83 users)

For the same 83 users, at MEILI there were 355 trips registered

On MEILI and PP (4 November, 83 users)

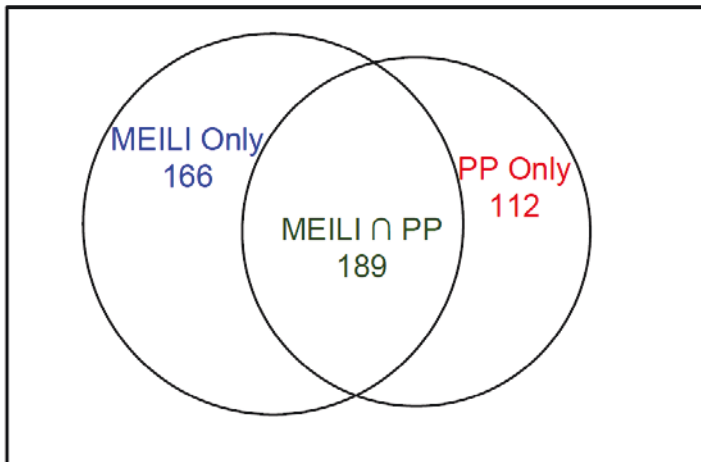




So what 2: Within this overlap day, how matched the trips were?

Trips are **matched** based on **temporal co-occurrence** (*start* and *stop* time of a MEILI trip has to be **within 30 minutes** of the *start* and *stop* time of a PP trip) and **identical purpose**

Total trips matched: 189



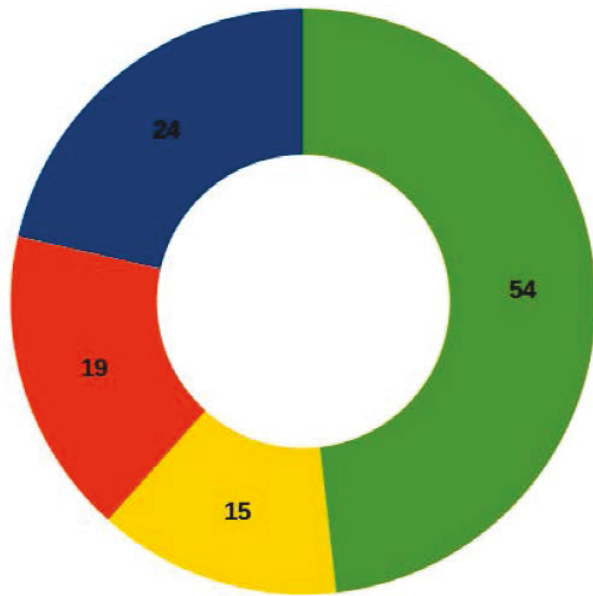
Missed movement / forgot smartphone:

MEILI did not collect either due to the minimum movement threshold (as determined by the accelerometer) or due to the fact that the user did not carry her smartphone

Purposes differ: trips that coincide but users declared different purposes

Trip chaining: within the period allocated for a trip in one system, there are more trips in the different system

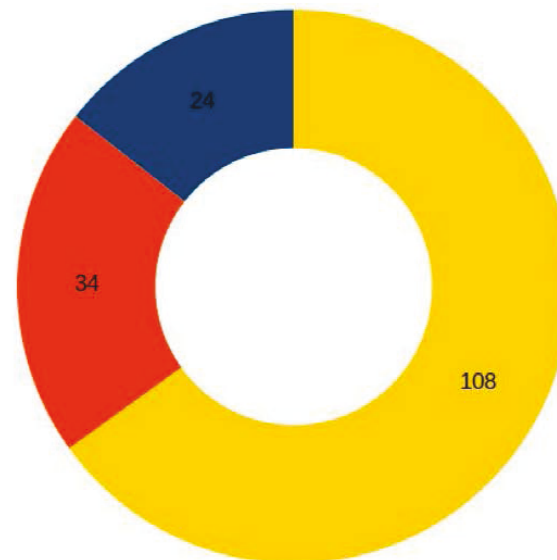
Missed by a system due to other reasons:
other reasons



MEILI Only: (N=166)

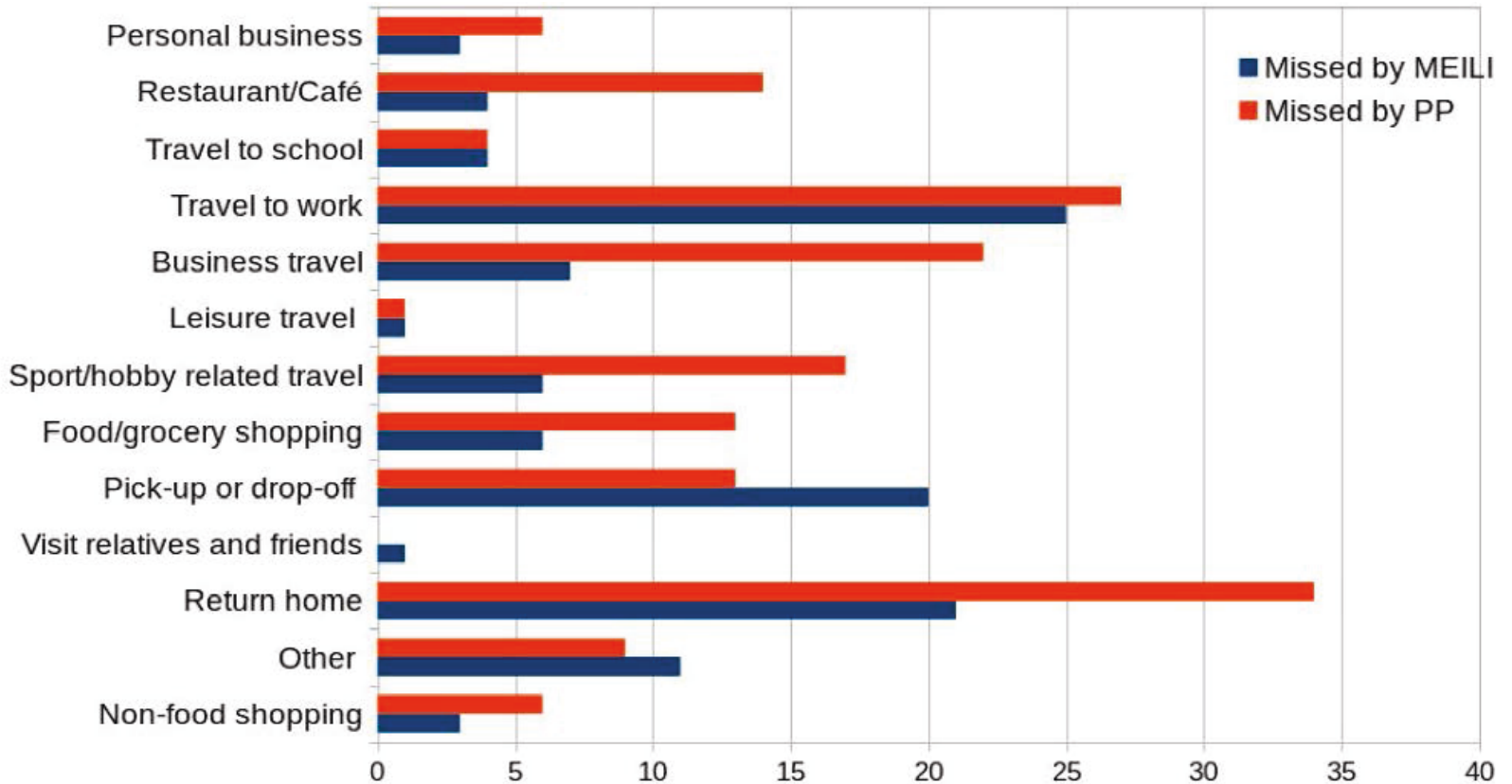
- Time coincides, purposes differ
- Trip chaining
- Missed by MEILI due to other reasons
- Missed movement

PP Only: (N=112)



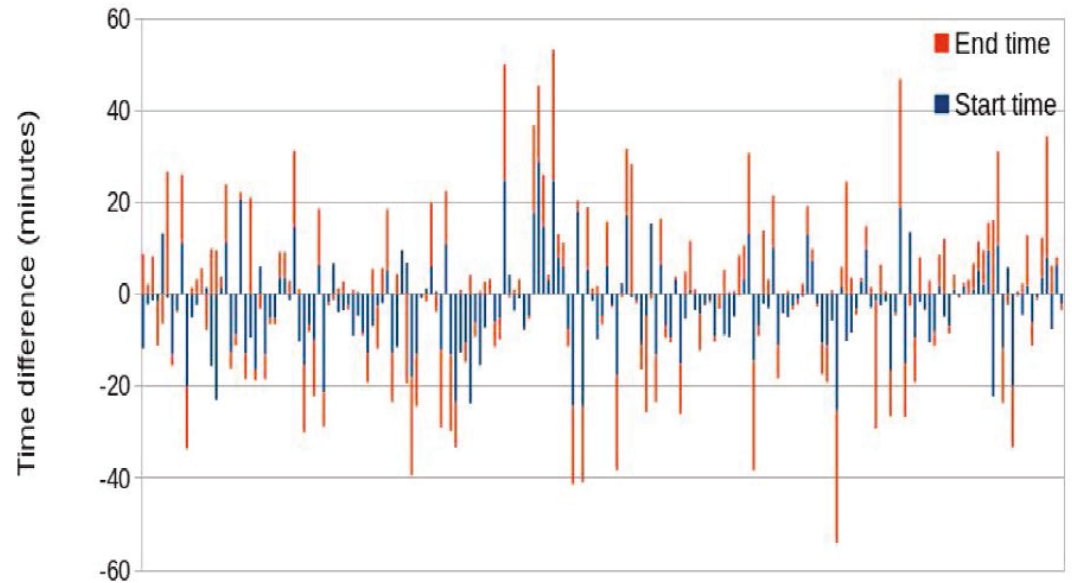
- Time coincides, purposes differ
- Trip chaining
- Missed by PP due to other reasons

Overview of the missed trips



Time estimation - comparison

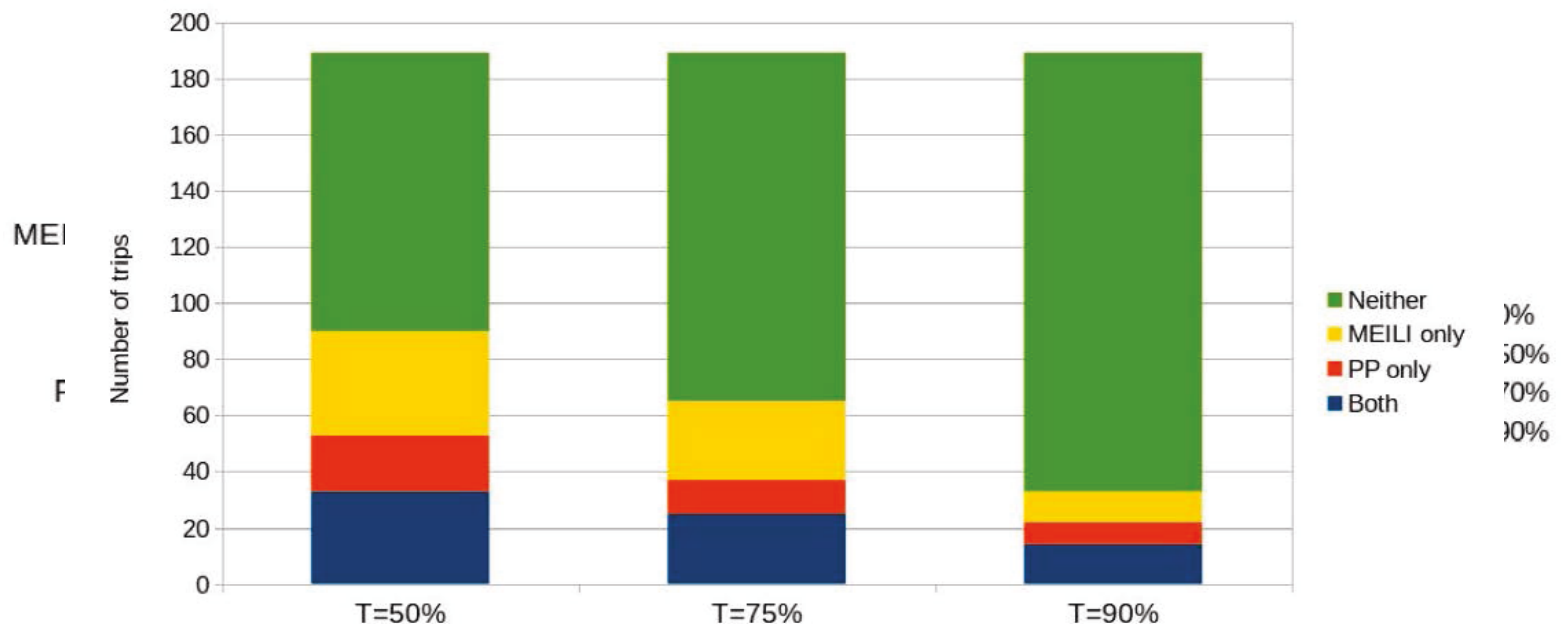
- The **difference** between the **start** and **end time** of a trip was 8 ± 7 mins
- The **maximum time difference** is **28 minutes**
- **Trend** in time difference: if the user **overestimates the start time**, he will **overestimate the end time** also.



So what 3: how to identify the ground truth and which one would perform better?

We introduced spatial-temporal indicators, (from 1% or very low spatial/temporal quality to 100% or very high spatial/temporal quality, route vs declared time interval).

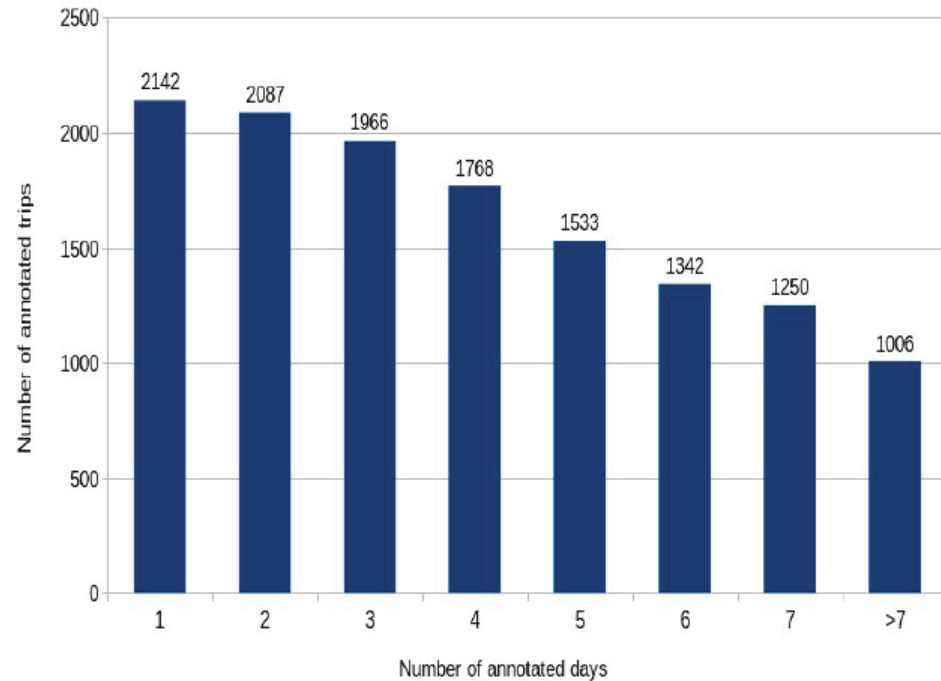
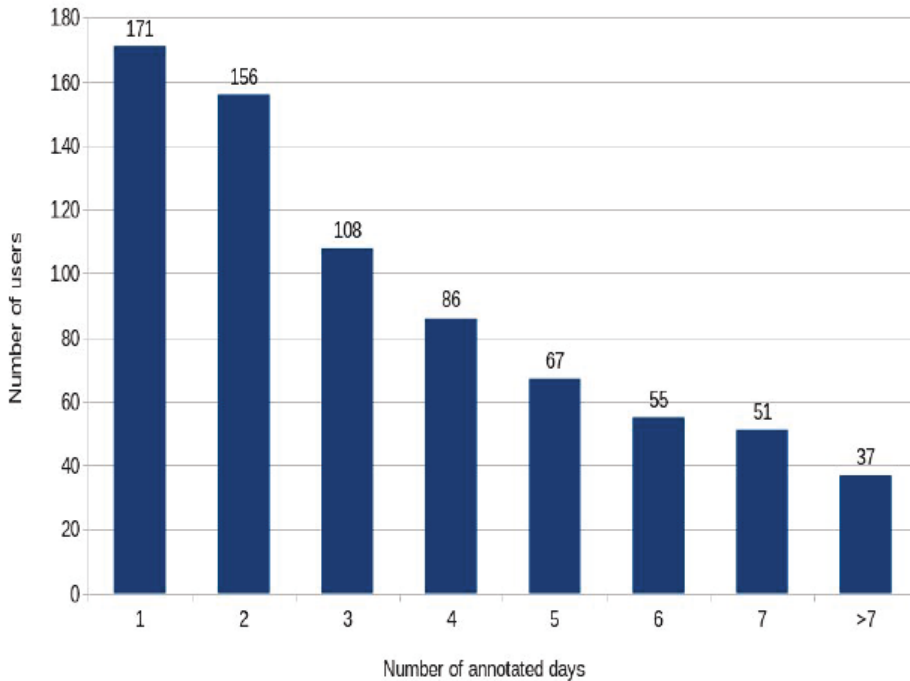
Res: 100 meter & 1 minute:



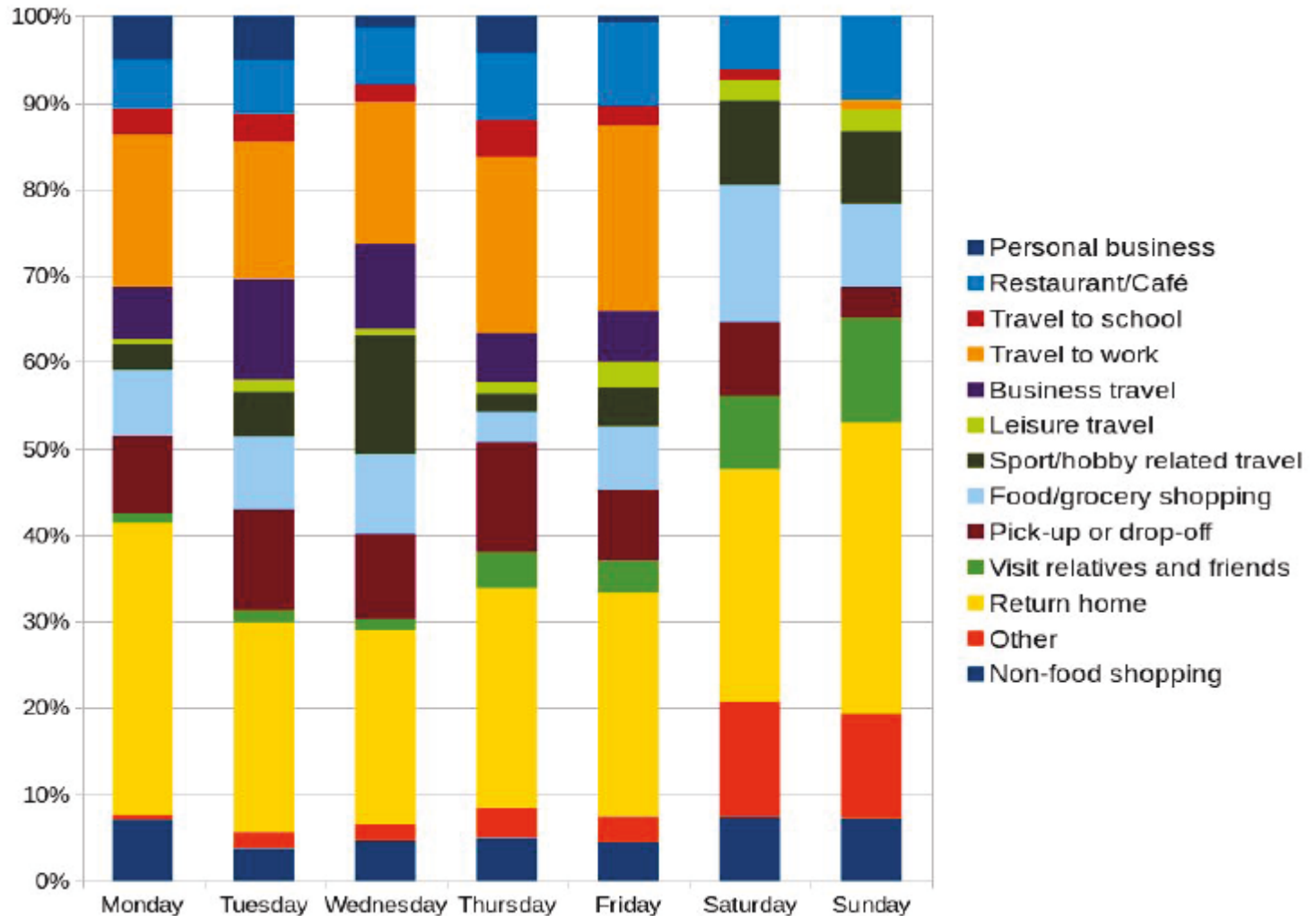
On average, MEILI has a lower error in time, but higher length error when making mistakes than PP



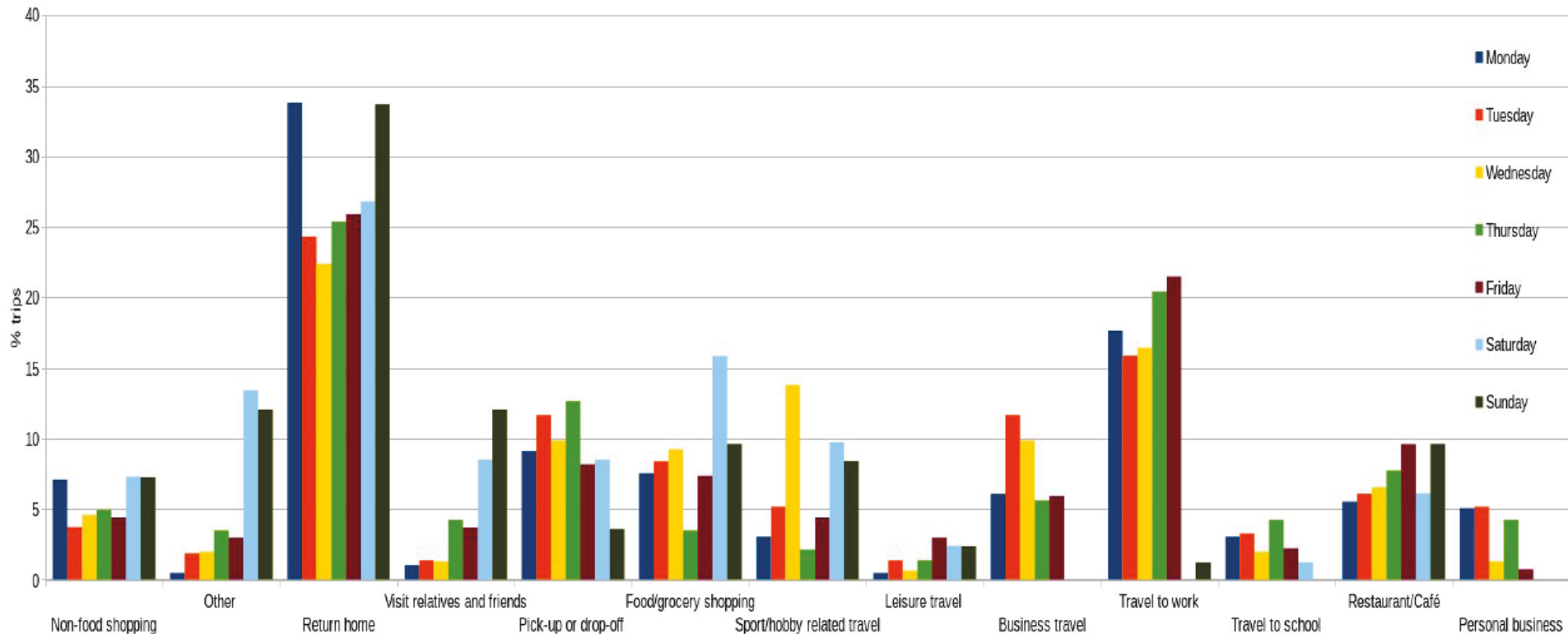
What's the profile with the rest of the MEILI (9 days) datasets?



So what 4: how stable the patterns were?



Daily variability by reported trip purposes





	# TI	Distance (km)		Duration(min)		S_idx(%)		T_idx (%)	
		Avg±Sd	Median	Avg±Sd	Median	Avg±Sd	Median	Avg±Sd	Median
Bicycle	346	3.1±8	0.5	194±641	20	82±27	99	51±39	83
Bus	512	6.9±18	2.4	64±229	13	73±34	92	59±36	81
Car as driver	986	10.1±17.2	4.1	77±246	14	77±31	93	62±35	81
Car as passenger	284	7.9±25	0.5	103±401	12	82±30	100	51±40	83
Commuter train	246	9.9±14.1	5.0	76±191	20.5	56±41	79	32±31	51
Ferryboat	70	2.5±17.1	0.1	42±80	17.5	84±32	100	8±15	17
Flight	5	601.4±385	444.3	94±36	80	1±2	0	30±46	21
Moped / Motorcycle	52	6±8.7	0.4	64±198	18.5	84±27	100	49±38	89



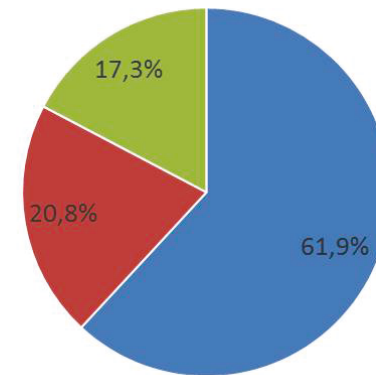
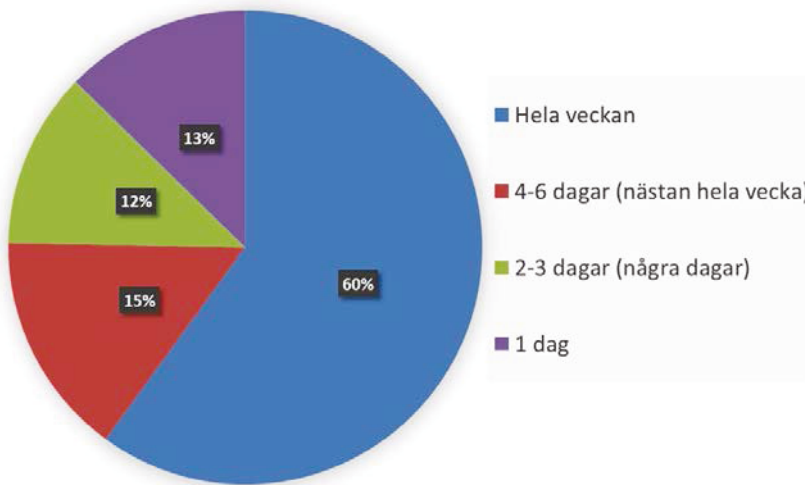
	# TI	Distance (km)		Duration(min)		Sp Ind (%)		T ind (%)	
		Avg±Sd	Median	Avg±Sd	Median	Avg±Sd	Median	Avg±Sd	Median
Other	8	70.3±178.9	0.4	1277±2842	235.0	75±35	91	9±8	7
Subway	431	9.8±69.2	4.1	53±180	19.0	37±41	55	23±28	32
Taxi	481	1.1±4.7	0.2	147±249	13.0	82±31	100	20±29	39
Train	137	16.7±83.8	0.2	526±940	427.0	78±36	100	18±28	55
Tram	253	1.7±12.4	0.1	58±145	12.0	80±34	100	24±32	96
Walk	2150	1.5±22.7	0.3	54±216	5.0	81±29	100	50±35	67
Total	5961					75±13	86	35±19	57

So what 5: what did the users think?

303 users that answered the feedback questionnaire
87% had no problem in installing the app.

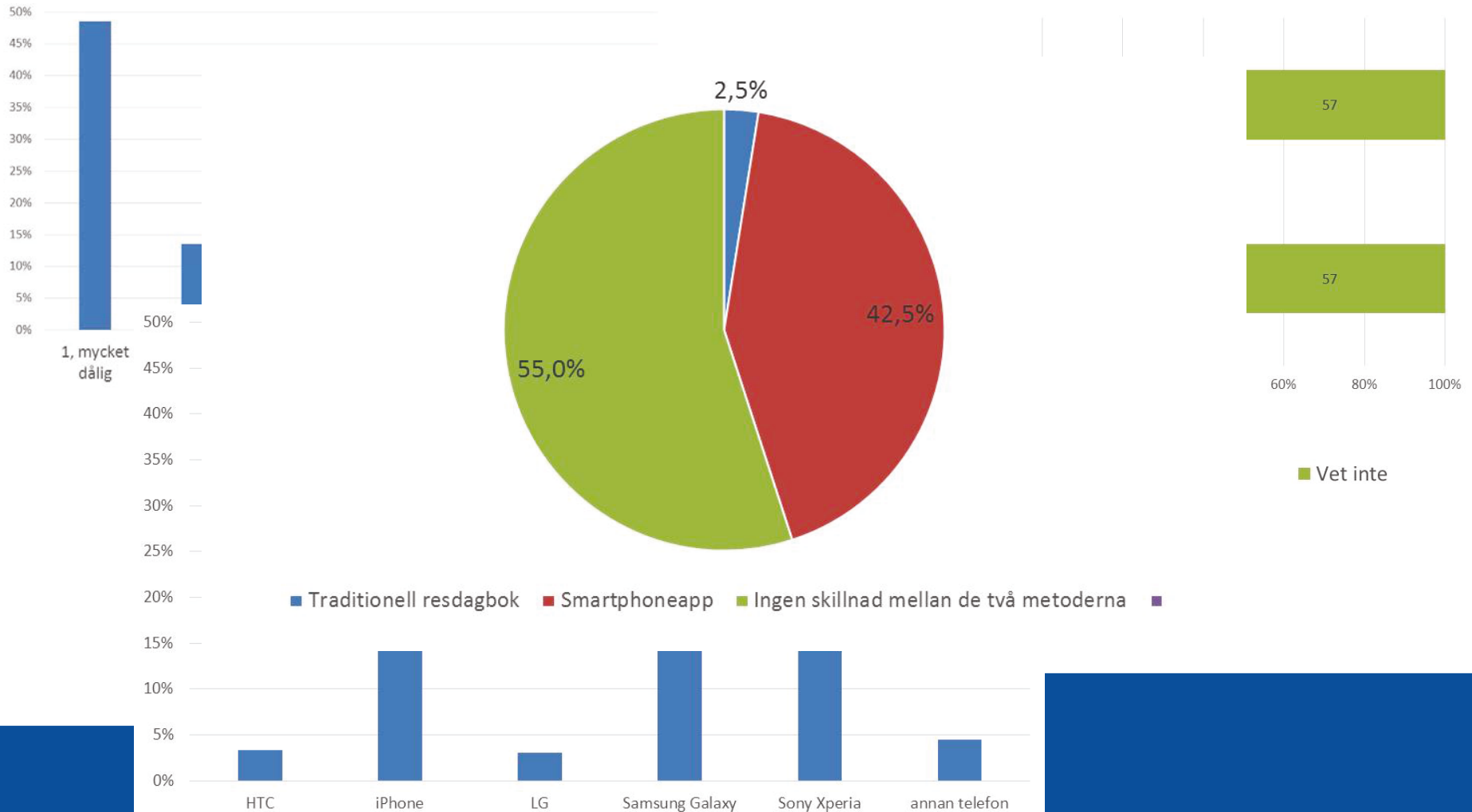
Most of them install the app for more than 3 days

65% of the ones who did not
installed the app said it doesn't work
Only 17% of them need to charge
the phone much often than usual



- Jag märkte ingen större skillnad mot normalt
- Jag behövde ladda telefonen lite oftare
- Jag behövde ladda telefonen mycket oftare

The web interface, however, were not well received





So what 6: What is needed for a successful application of a survey app like Meili?

MEILI as a flexible, transferable, open, data collector does work nicely – technology wise, **doable and it works**.

App is not "just a survey tool", but is seen as a product. Needs more complete multidisciplinary team (design, bugs, system/survey support, communication with users)

As launching a product, we need to understand the platforms better (e.g. contact with Apple, understand limitations of different browsers)

Also know how to "manage the users" since this is not just a one-day person trips, but multi-day observations (more intensive interaction with users)

Bias in recruitment? Or more like identifying markets and treating it as a complementary system? Or deploy different strategy in recruiting/selling it to different groups of users (stratified sampling)



What's next?

1. How far such big data is actually usable in our modelling approach:

- The stability and variability of individuals' day-to-day choices
- The comparability of the current national transport model indicators and assumptions with GPS based observations
- Simulating individual space-time prisms

2. Small trial in Gothenburg (November 2016)



References:

Source code for MEILI <https://github.com/Badger-MEILI>

Mobility Collector - Prelipcean, A. C., Gidófalvi, G., & Susilo, Y. O. (2014). Mobility collector. *Journal of Location Based Services*, 8(4), 229-255.

A framework for the comparison of travel diary collection systems - Prelipcean, A. C., Gidófalvi, G., & Susilo, Y. O. (2015).

Comparative framework for activity-travel diary collection systems. In *Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2015 International Conference on* (pp. 251-258). IEEE.

On **AI performance measures** relevant to travel diaries - Prelipcean, A. C., Gidófalvi, G., & Susilo, Y. O. (2016). Measures of transport mode segmentation of trajectories. *International Journal of Geographical Information Science*, 30(9), 1763-1784.

Susilo, Y.O., Prelipcean, A.C., Gidófalvi, G., Allström, A., Kristoffersson, I., and Widell, J. (2016) **Lessons from a trial of MEILI**, a smartphone based semi-automatic activity-travel diary collector, in *Stockholm city, Sweden. World Conference on Transport Research 2016, Shanghai. China.*



THANK YOU

QUESTIONS?





And then ...

1. Machine learning algorithms are used to train classifiers that infer the information specified by the user (trip's destination, purpose, triplets and means of transportation)
2. By analyzing the accuracy of the classifiers, it can be established what parts of the traditional travel surveys can be automated
3. The traditional travel surveys and the new method that uses the MEILI system are compared, their advantages and disadvantages are discussed and a conclusion is drawn.



Large field trial (3-9 Nov) – Dates and facts

28 Oct	Information e-mail
2 Nov	E-mail about installing app and participation not possible for iPhone users
2 Nov at 10PM	App accepted and published on Appstore
3 Nov	New information to Iphone users
5 Nov	Paper&Pencil survey for 4nov
9 Nov	Information about incentives
11 Nov	Follow-up survey
16 Nov	Information about winners



Large field trial – Recruitment

21 Sept – 25 oct Users signing up/recruited mainly from travel survey in Stockholm

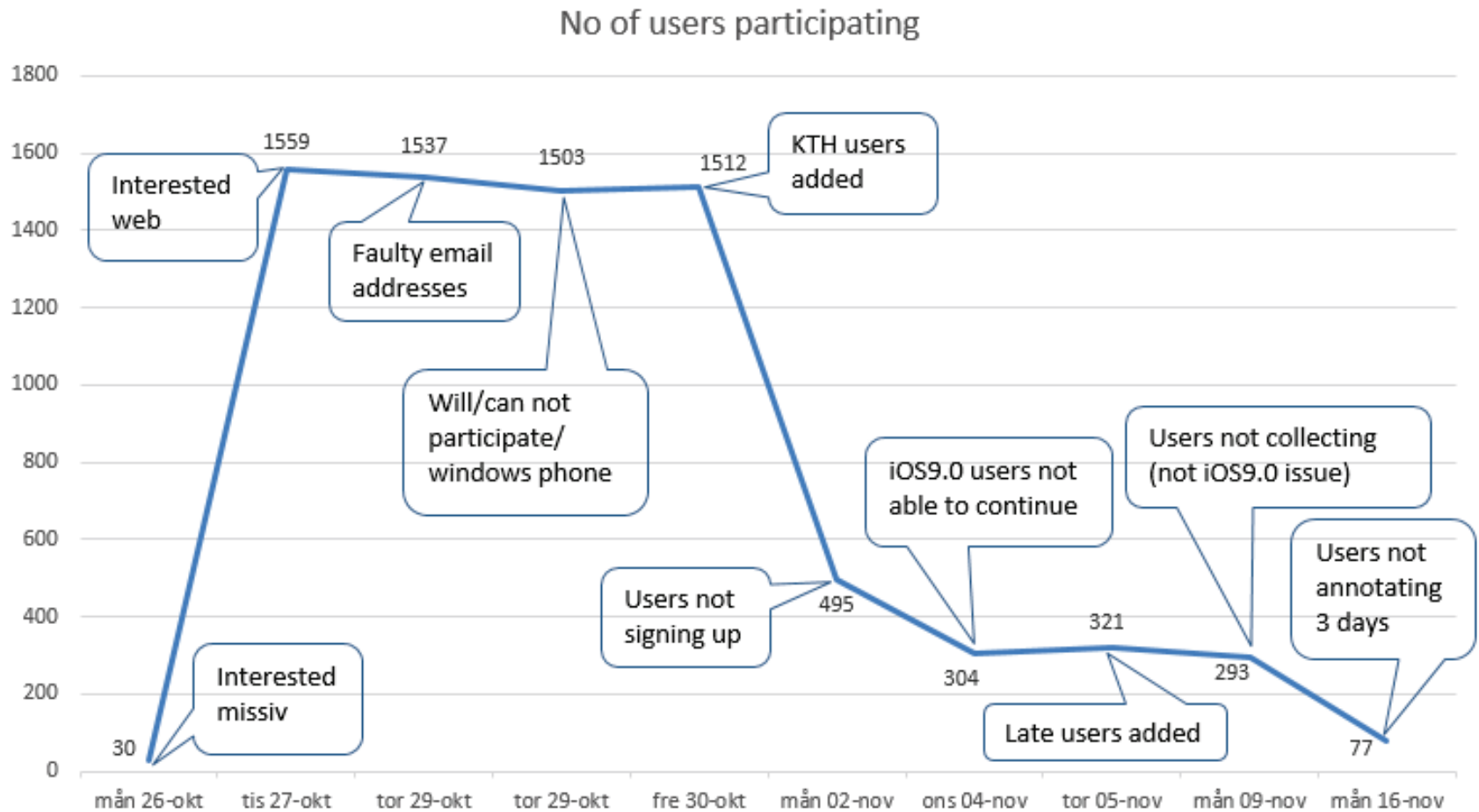
Missiv letter:

“Vill du delta i ett forskningsprojekt där du använder din smartphone för att samla in resdata? Projektet finansieras av Trafikverket och syftar till att förbättra metoderna för att samla in resdata, data som används som underlag vid planeringen av framtidens samhälle. Om du är intresserad - skicka ett mail till spot@sweco.se så får du mer information.”

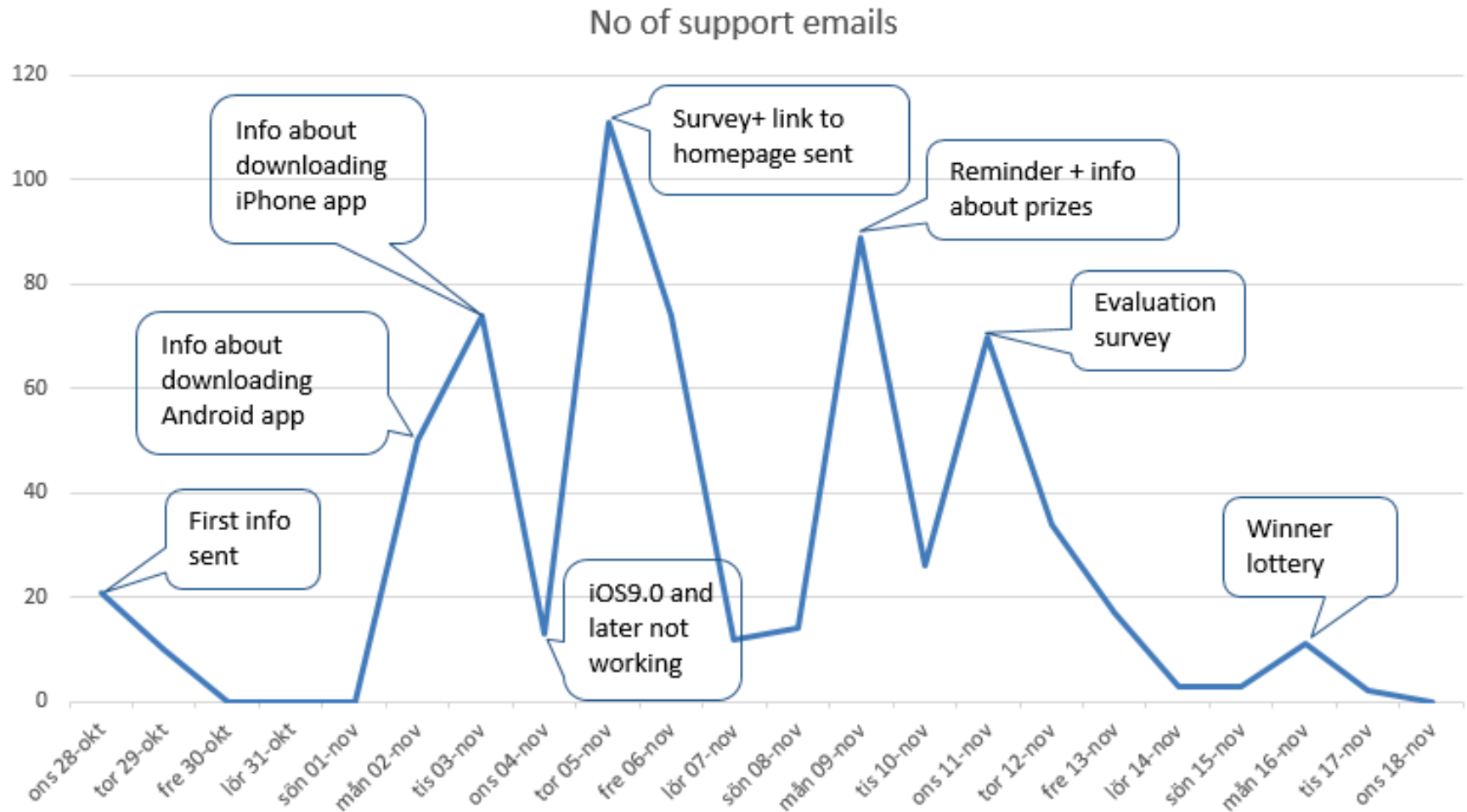
Web based survey as above and English version:

“Do you want to participate in a research project where you would use your smart phone to collect travel data? The project is financed by the Swedish Transport Administration and aims to improve the methods by which we gather travel data. If you are interested, please enter you e-mail below and we will send you more information.”

Number of participants during the field trial



Number of support e-mails





So what 2: What's the common questions?

Support

- Need new username and/or password
- No possibility to use Chrome
- Problem to log in/nothing is shown when logging in on webpage
- Is stuck on a trip and cannot move on to annotate more trips
- "Invalid state" because of too many clicks on webpage

Feedback

- Questions about bias and the project in general
- Constructive comments about how the webpage can be improved
- Complaints about the webpage



Paper&pencil – group statistics

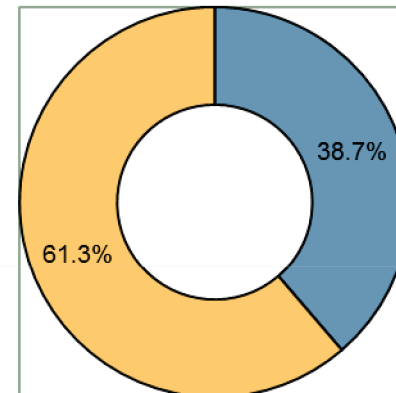
Web-based

Sent out before the users started annotating their data

Sent out to 510 users 431 answered → 86%

16 to 77 years old, average age 42

All living in the Stockholm area



■ Kvinna
■ Man



Data definition

Criteria for choosing a user for the study set:

- Email addresses recorded by both systems have to **match exactly** (to avoid partial string matching)
- At least one annotated trip has to be partially within **the comparison day** (4 Nov)

***Obs!* Not an exhaustive user set extraction due to time limitations**

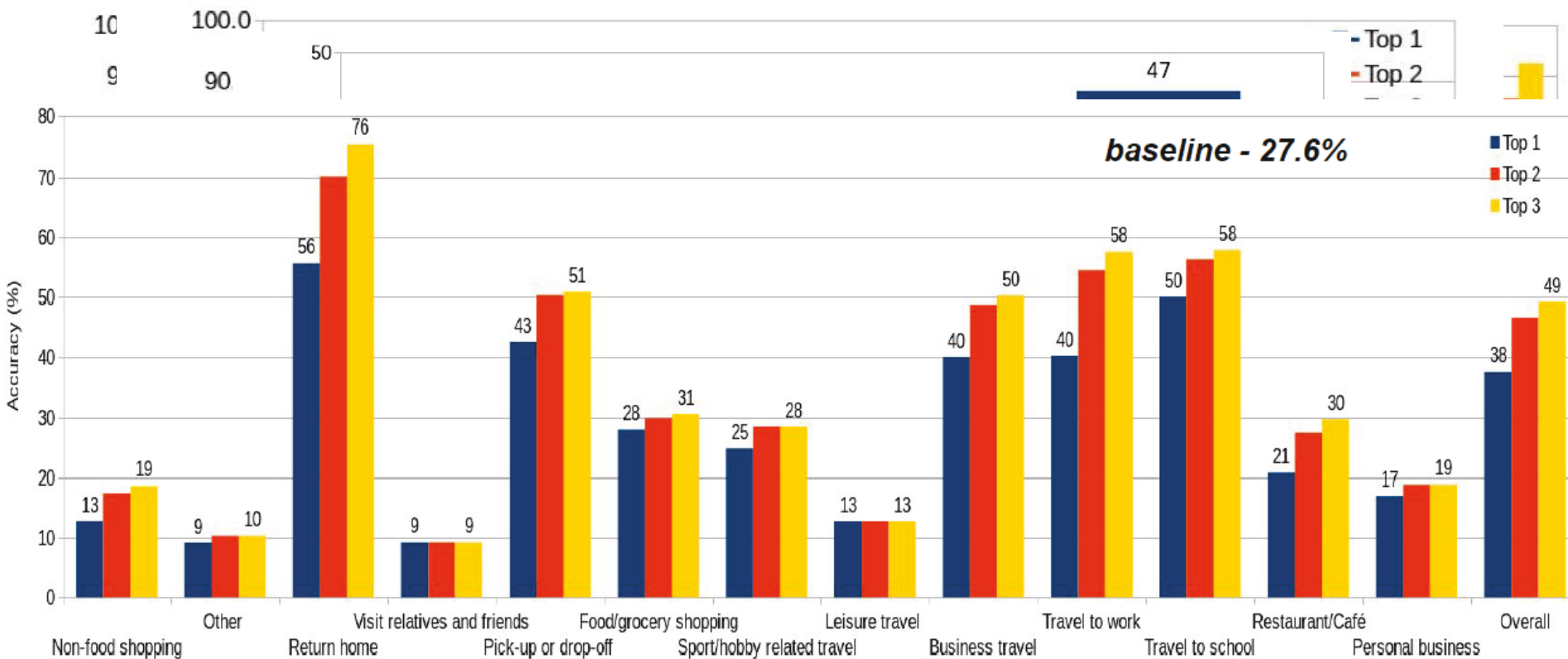


Accuracy of inference methods

Stop detection: 96.7%

Trip-leg: 70.2%

Per trip: 78.5%





Summary and conclusions

This project aims to: “refine, test and deploy a prototype of a smartphone application ...” (24 months)

We were hoping to get 300, 1559 expressed interest, 495 signed up, but only 293 participated, 171 annotated at least one day and 77 annotated 3 days and so on.

Original plan was only Android, but managed to prepare the iPhone version as well, before had problem with the Apple’s bureaucracy (apparently being seamless/non-intrusive and non-business oriented is not a good thing)

2142 annotated trips from 171 users (1250 trips, from 51 users for a duration longer than a week) collected in field trial, can be said as one the most successful trials in the world

About 1 million GPS readings from 293 users.

Web page for annotation most critical