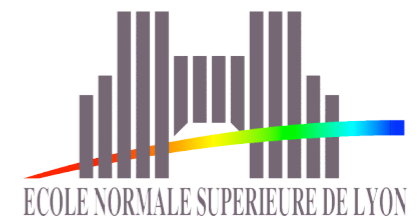


Patient - Healthcare Workers interactions using a WSN

Adrien Friggeri <adrien.friggeri@ens-lyon.fr>
Guillaume Chelius <guillaume.chelius@inria.fr>
Eric Fleury <eric.fleury@ens-lyon.fr>



raw data

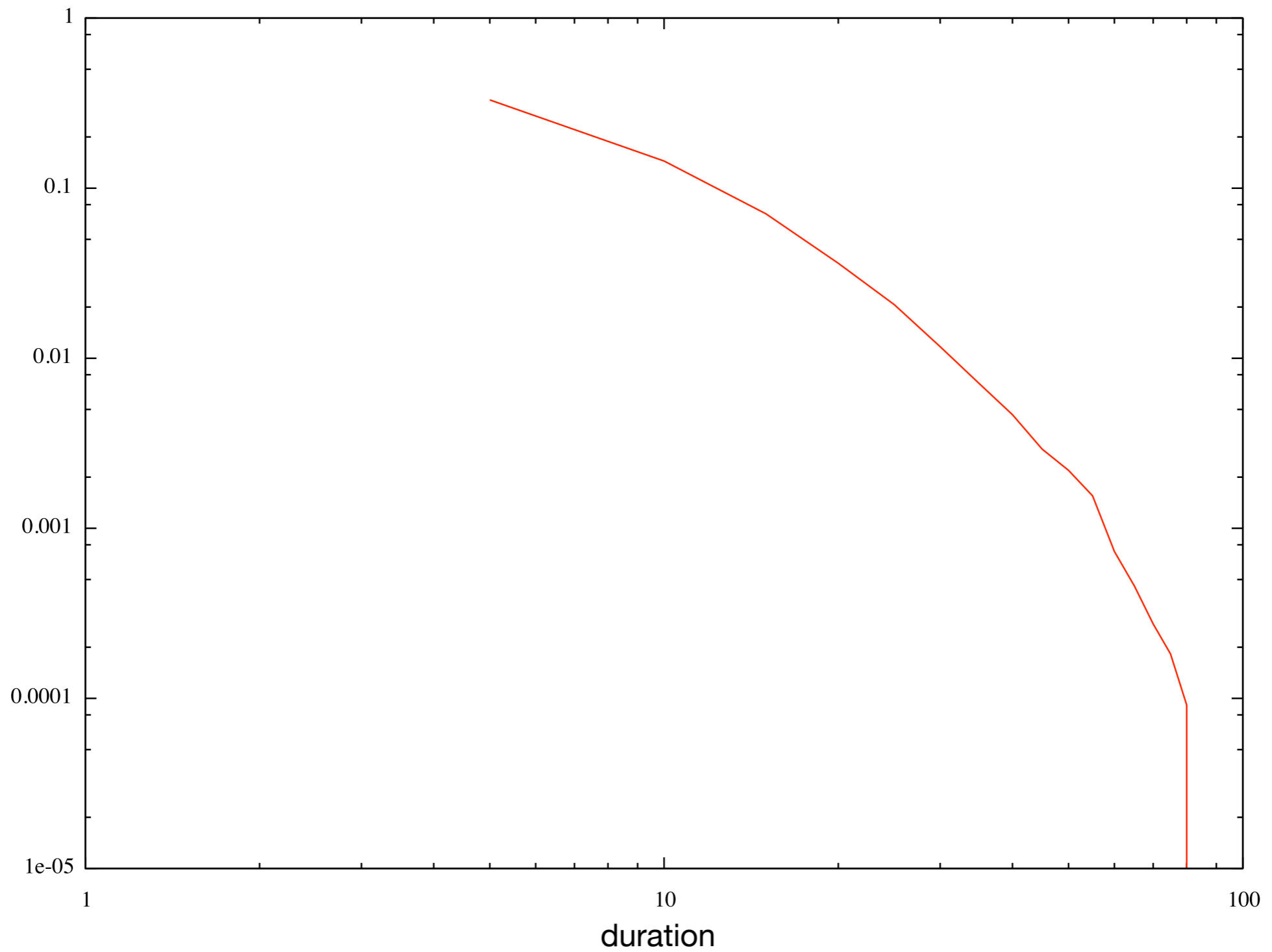
- ▶ 98 days
- ▶ 16,066,096 contacts
- ▶ 56 mobile sensors
- ▶ 32 fixed sensors (rooms)

raw data

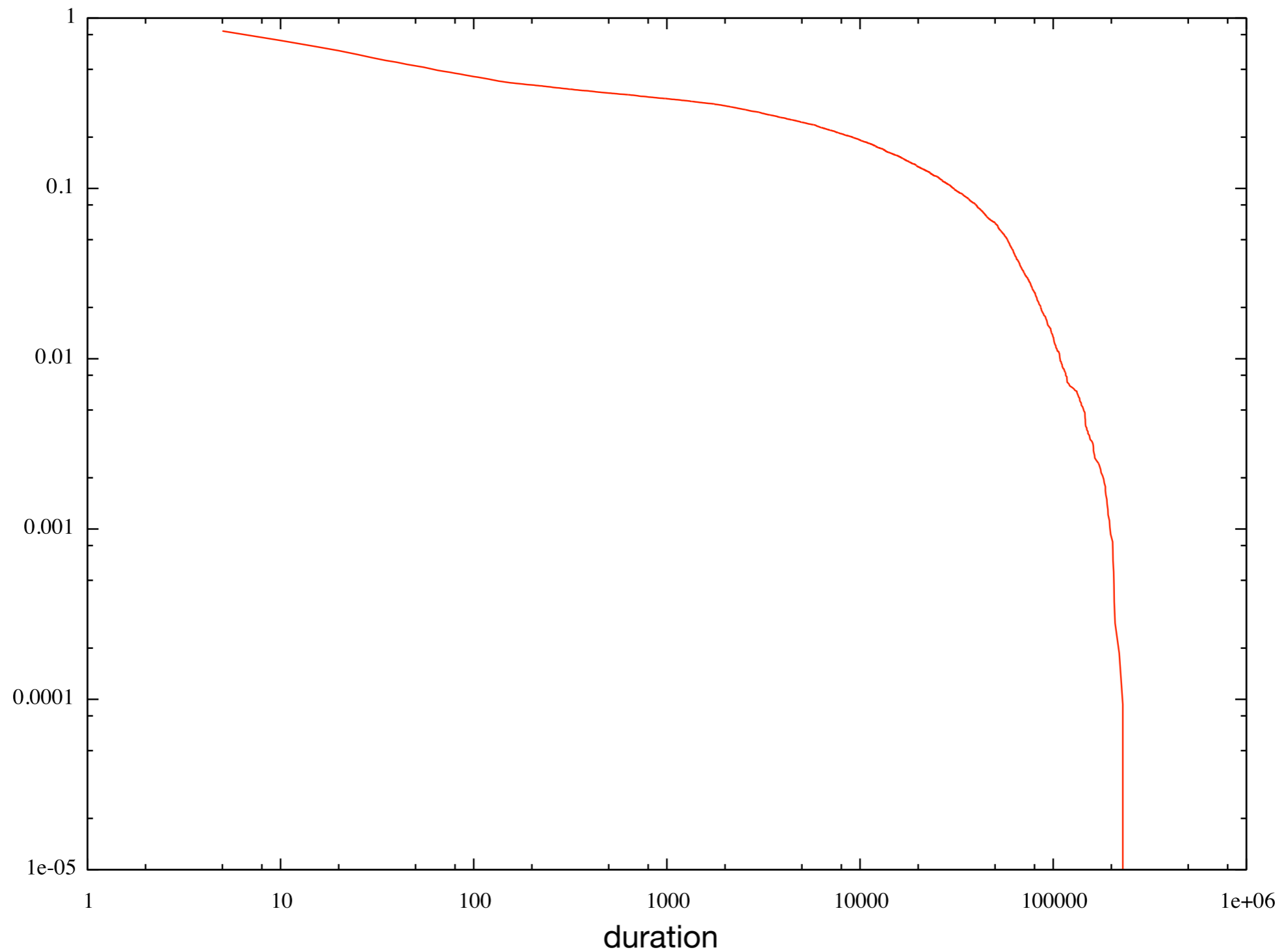
```
mysql> select time, fixed, mobile, level from contacts limit 10;
+-----+-----+-----+-----+
| time      | fixed | mobile | level |
+-----+-----+-----+-----+
| 1231196400 | 6     | 23     | 1     |
| 1231196400 | 21    | 14     | 1     |
| 1231196400 | 21    | 43     | 2     |
| 1231196400 | 21    | 6      | 1     |
| 1231196400 | 16    | 6      | 1     |
| 1231196400 | 16    | 9      | 1     |
| 1231196400 | 23    | 43     | 1     |
| 1231196400 | 23    | 33     | 1     |
| 1231196400 | 8     | 32     | 1     |
| 1231196400 | 17    | 5      | 1     |
+-----+-----+-----+-----+
10 rows in set (0.00 sec)
```

- ▶ 1 contact :
 - ▶ timestamp
 - ▶ couple of sensors
 - ▶ RSSI (1 : weakest, 4 strongest)

durations c.d.f



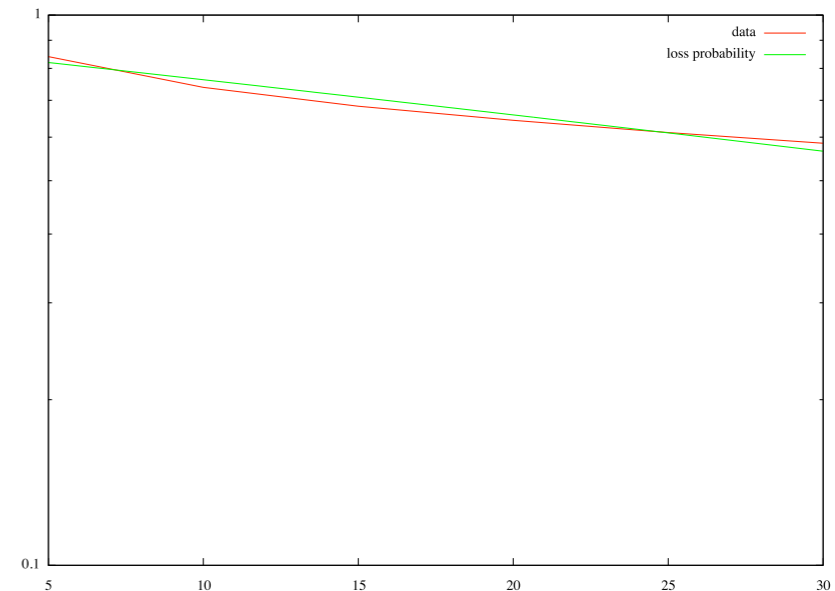
intercontacts c.d.f



packet loss

Presence of short intercontacts can be explained by packet loss :

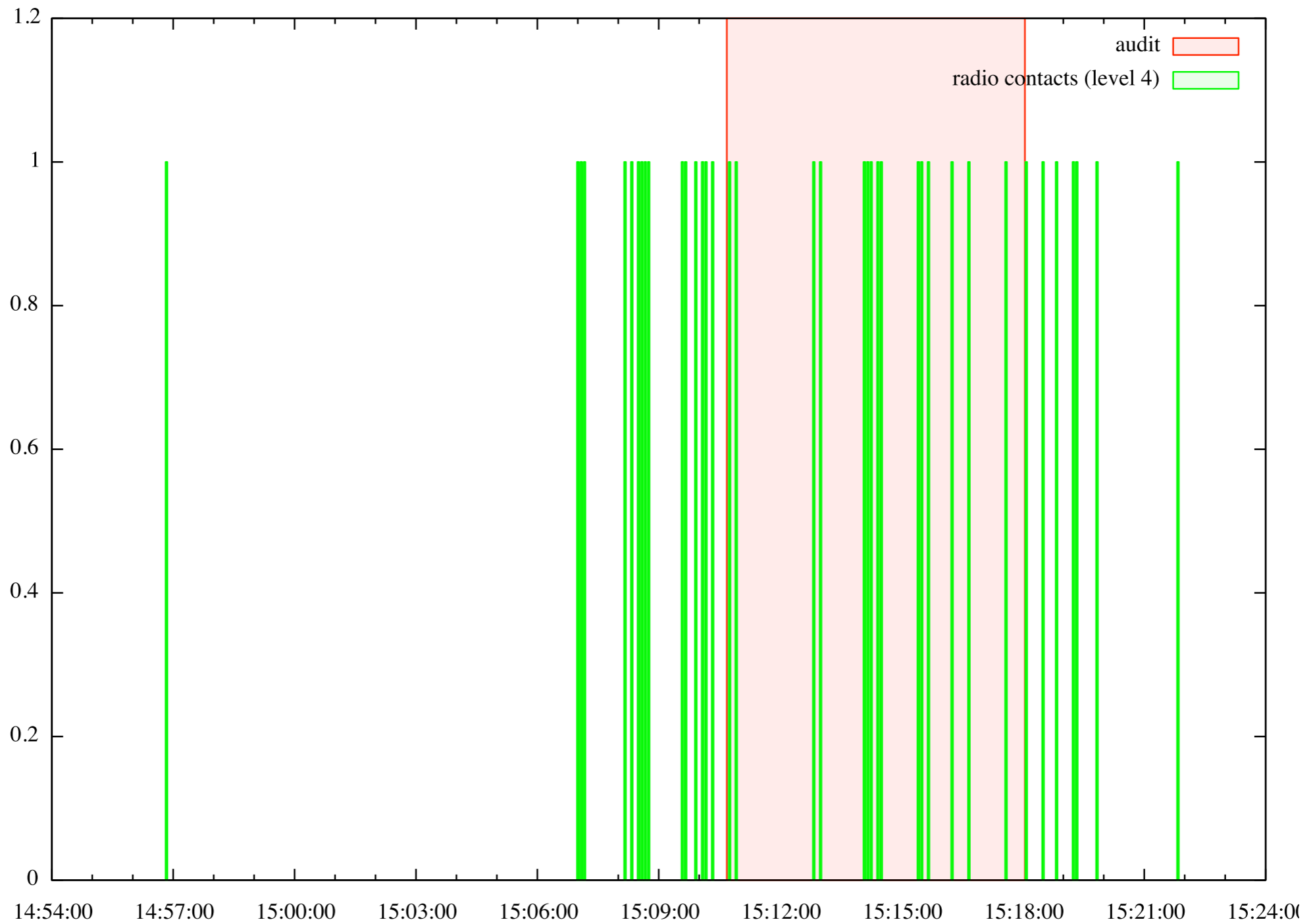
- ▶ let p_4 be the probability of receiving a packet when there is a presence
- ▶ the probability of having an intercontact of $5 \cdot k$ seconds for small k is $(1-p_4)^{-k}$
- ▶ $p_4 \sim 0.13$



Several visits where measured **by hand**.

- ▶ time of entry
- ▶ duration of the visit
- ▶ sensors ids

audits



using audits to determine p_4

$$p_4 \sim \frac{5N_{\text{contacts}}}{T_{\text{audits}}}$$

- ▶ $p_4 \sim 0.145$
- ▶ consistent with value obtained using intercontacts

isolation triplets

Question :

Do rooms in isolation receive a different treatment ?

isolation triplets

Question :

Do rooms in isolation receive a different treatment ?

Data :

- ▶ Series of triplets of rooms
 - ▶ 1 isolated room
 - ▶ 2 control rooms
 - ▶ beginning & end date of inclusion (3 days)

isolation triplets

15 triplets :

- ▶ 1 without data (failing sensors)
- ▶ 2 with abnormal data (including room adjacent to where sensors where stored)

isolation triplets

15 triplets :

- ▶ 1 without data (failing sensors)
- ▶ 2 with abnormal data (including room adjacent to where sensors were stored)

We focus on the 12 remaining triplets.

presence detection

2 metrics to characterize the treatment of a room :

- ▶ Number of **visits**
- ▶ Total **duration**

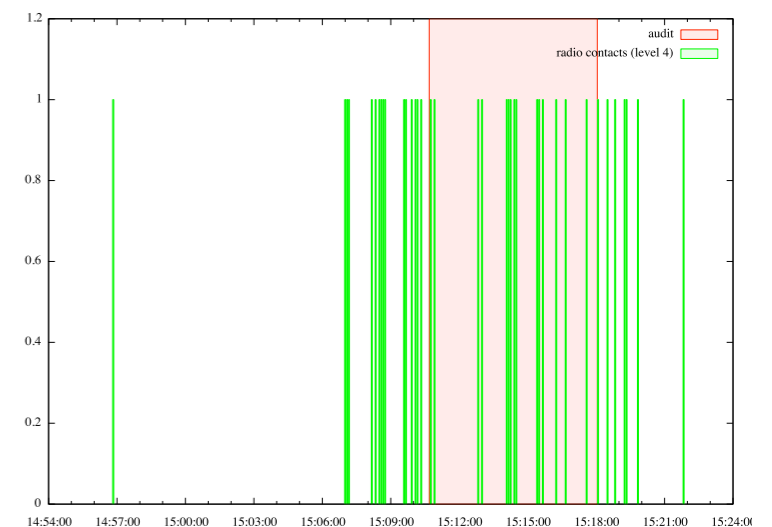
presence detection

2 metrics to characterize the treatment of a room :

- ▶ Number of **visits**
- ▶ Total **duration**

Considering only *real* contacts leads to :

- ▶ overestimating **visits**
- ▶ underestimating **duration**



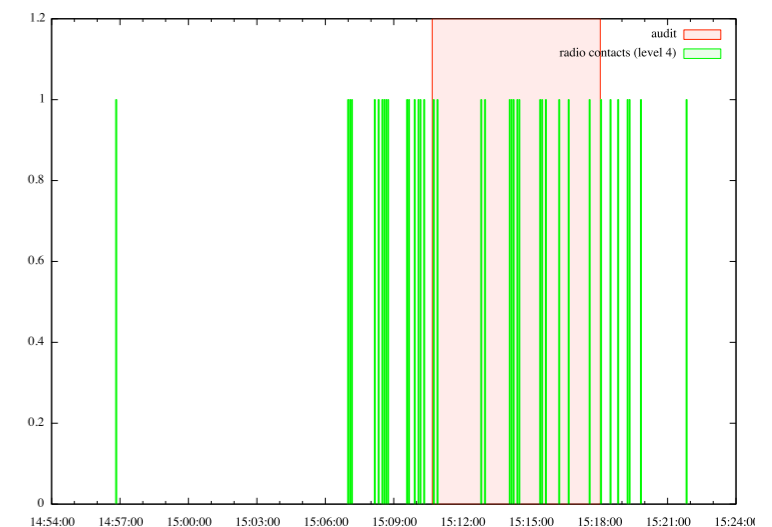
presence detection

2 metrics to characterize the treatment of a room :

- ▶ Number of **visits**
- ▶ Total **duration**

Considering only *real* contacts leads to :

- ▶ overestimating **visits**
- ▶ underestimating **duration**



Packet loss → need to reconstruct the presence signal

probabilistic reconstruction

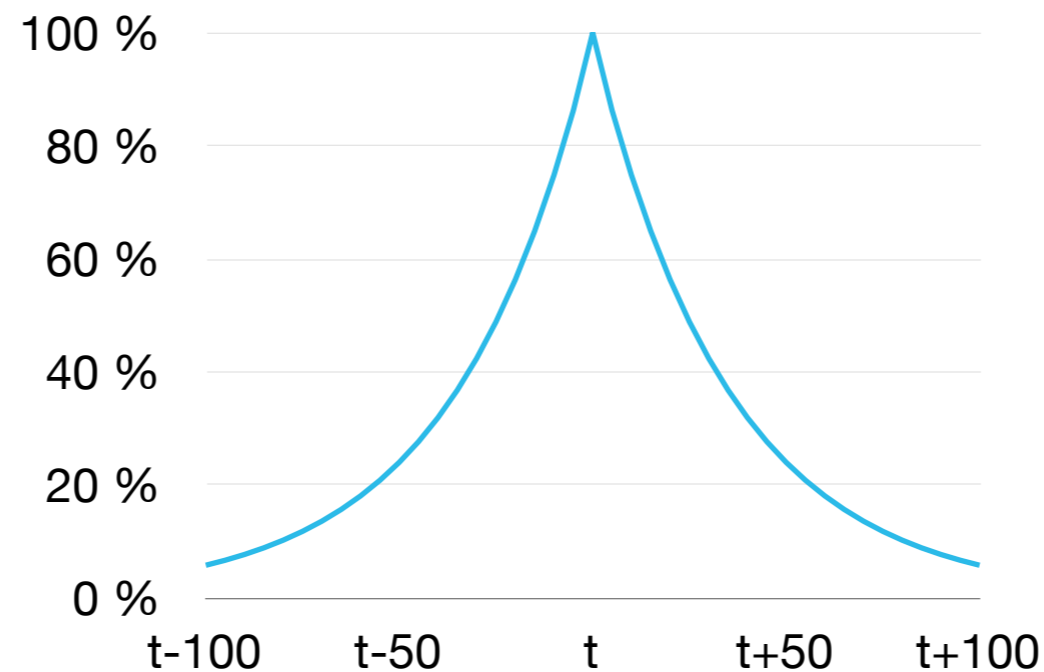
Suppose a packet is received at t .

probabilistic reconstruction

Suppose a packet is received at t .

- ▶ induces a probability of presence around t

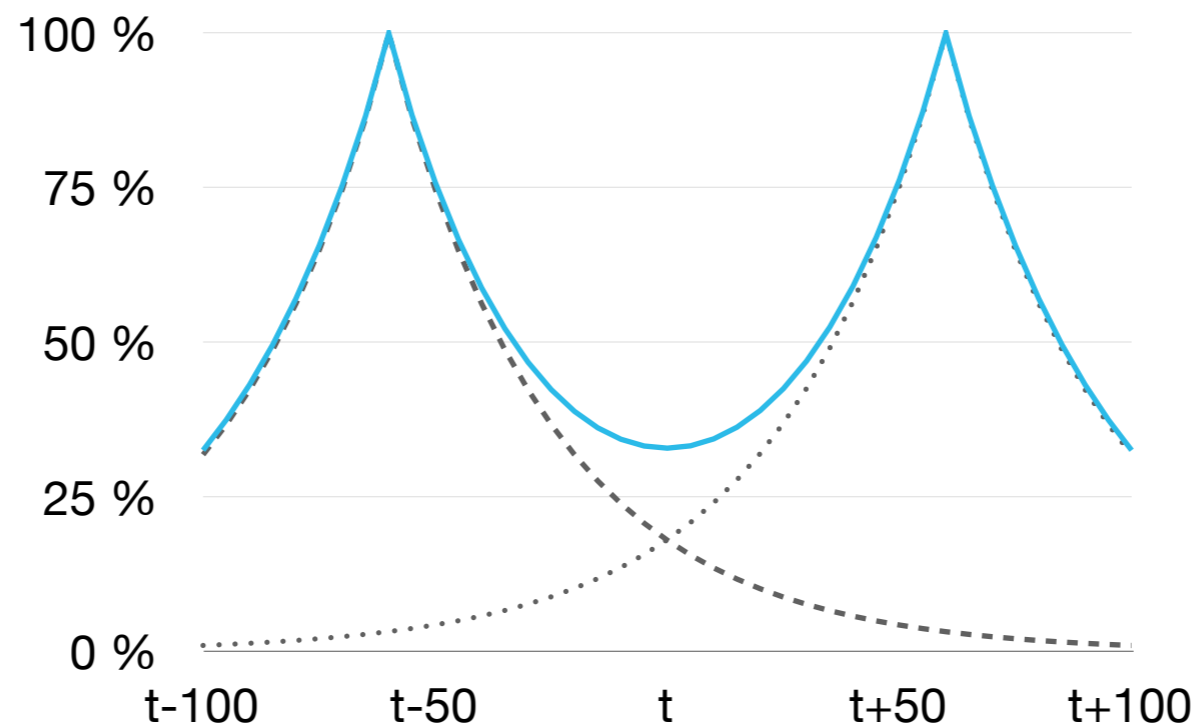
$$\mathcal{P}(\text{present at } t' | \text{seen at } t) = (1 - p_4)^{\frac{|t-t'|}{5}}$$



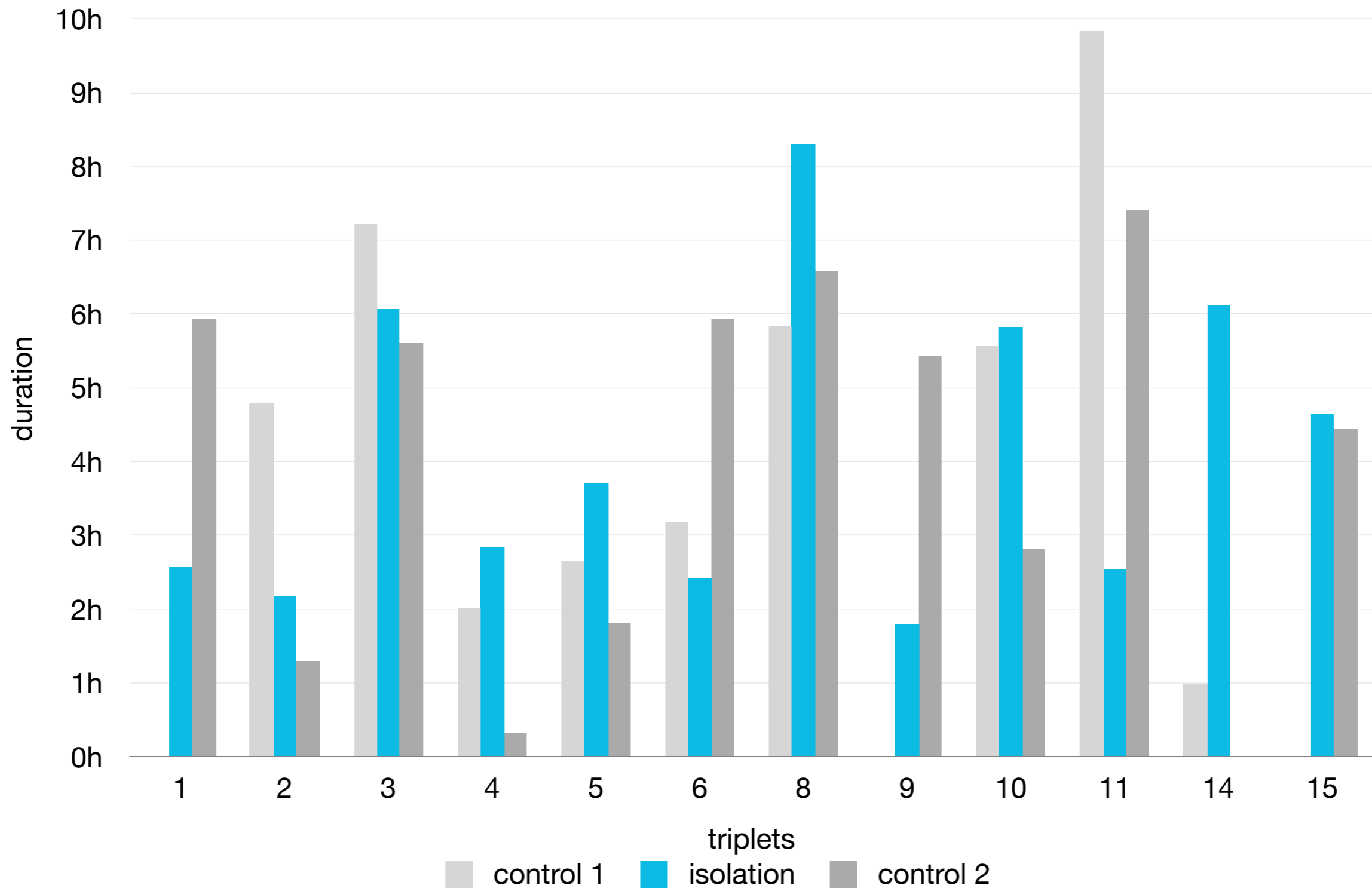
probabilistic reconstruction

Which can be generalized to:

$$\mathcal{P}(\text{present at } t' | \text{seen at } (t_i)) = 1 - \prod_{t_i} \left(1 - (1 - p_4)^{\frac{|t' - t_i|}{5}} \right)$$



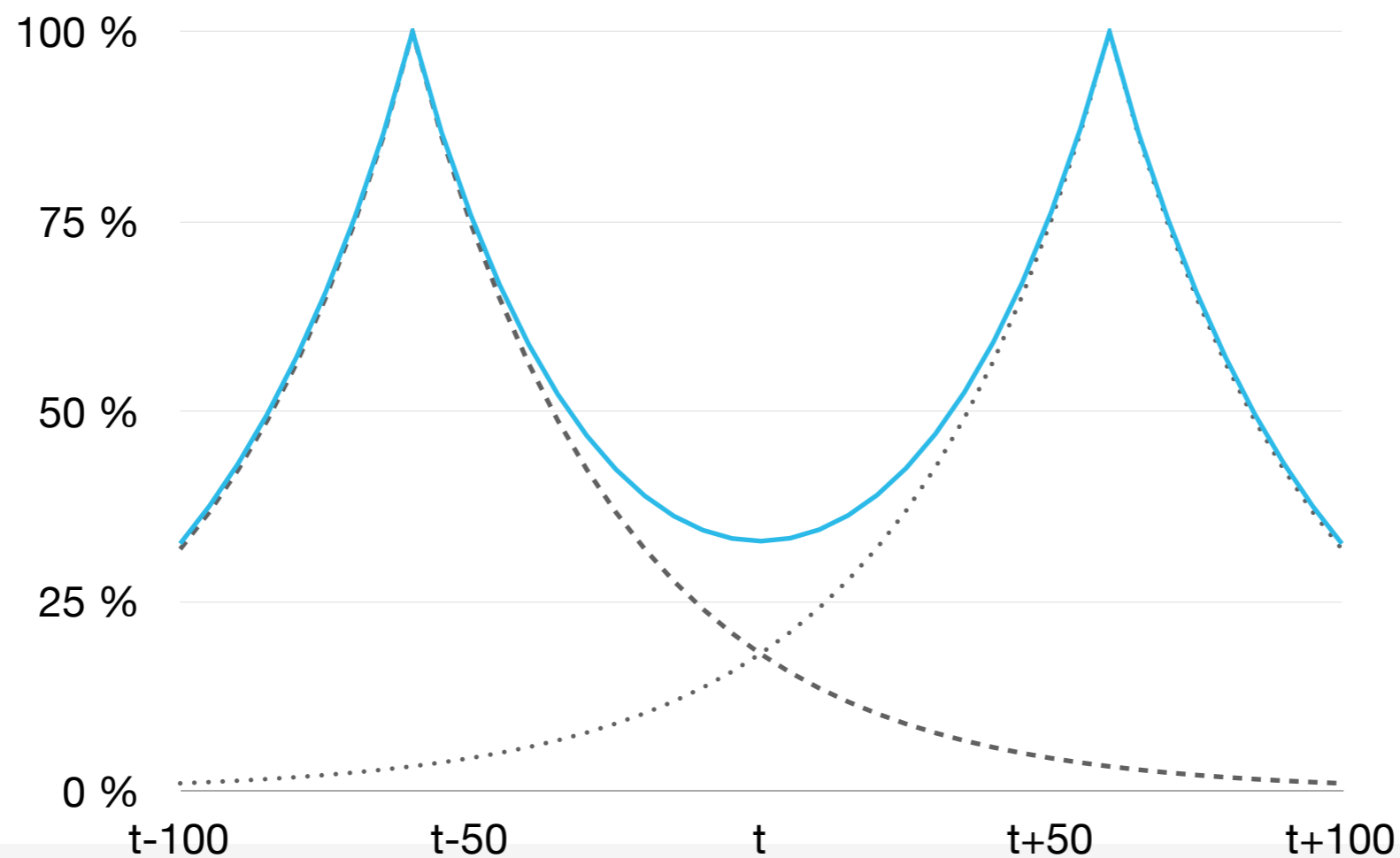
probabilistic reconstruction



drawbacks

Two problems :

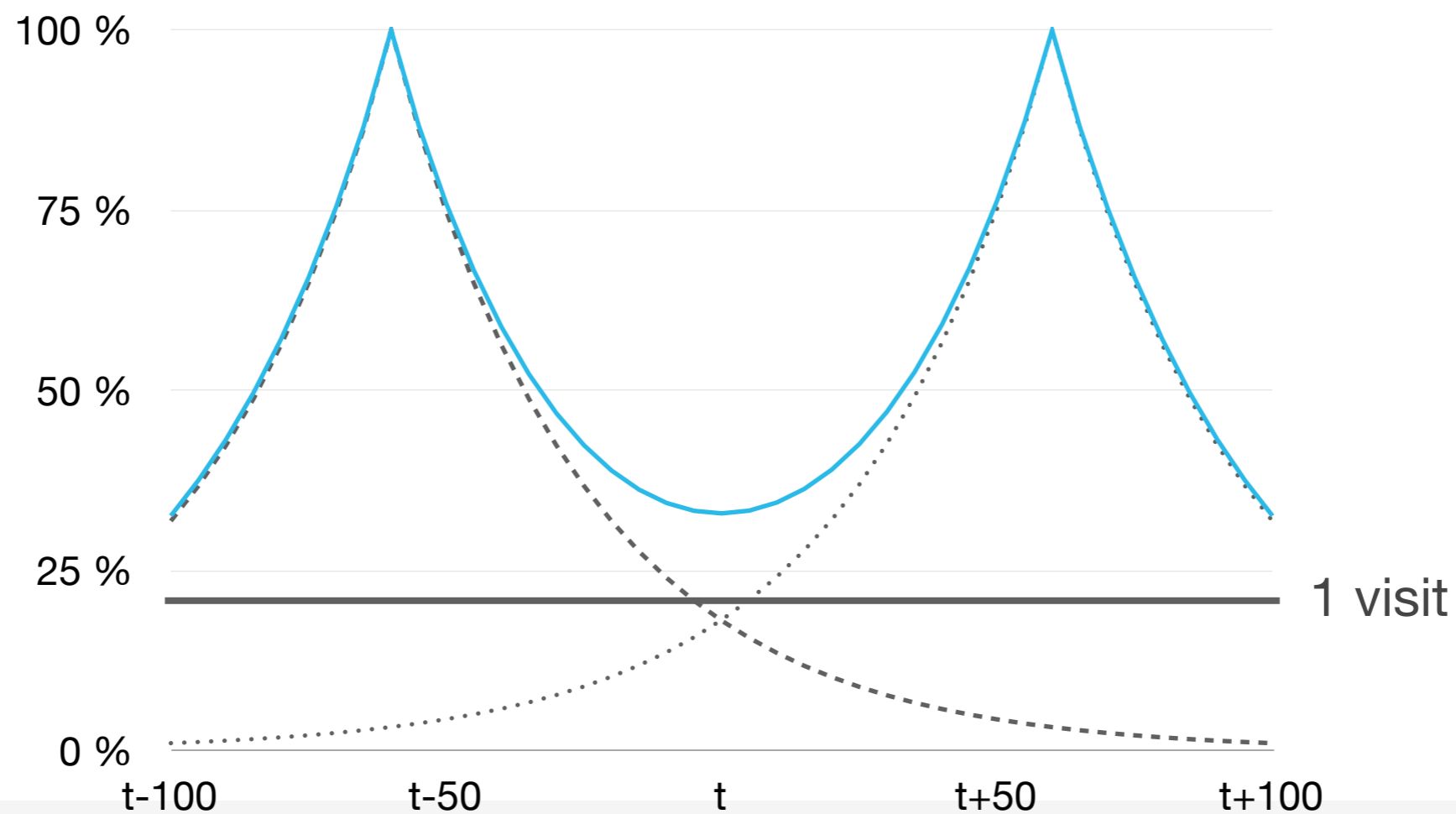
- ▶ Defining visits : threshold ?



drawbacks

Two problems :

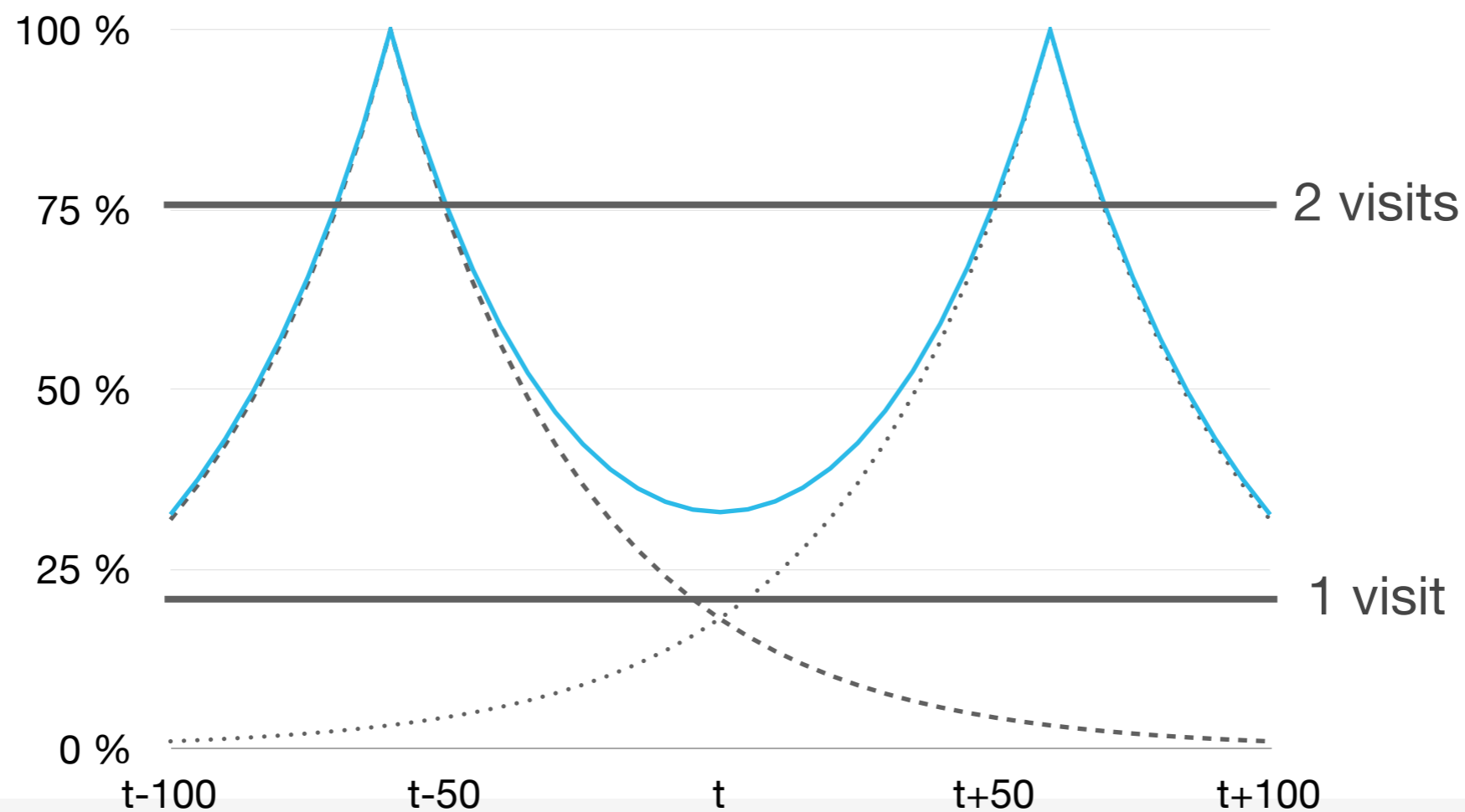
- ▶ Defining visits : threshold ?



drawbacks

Two problems :

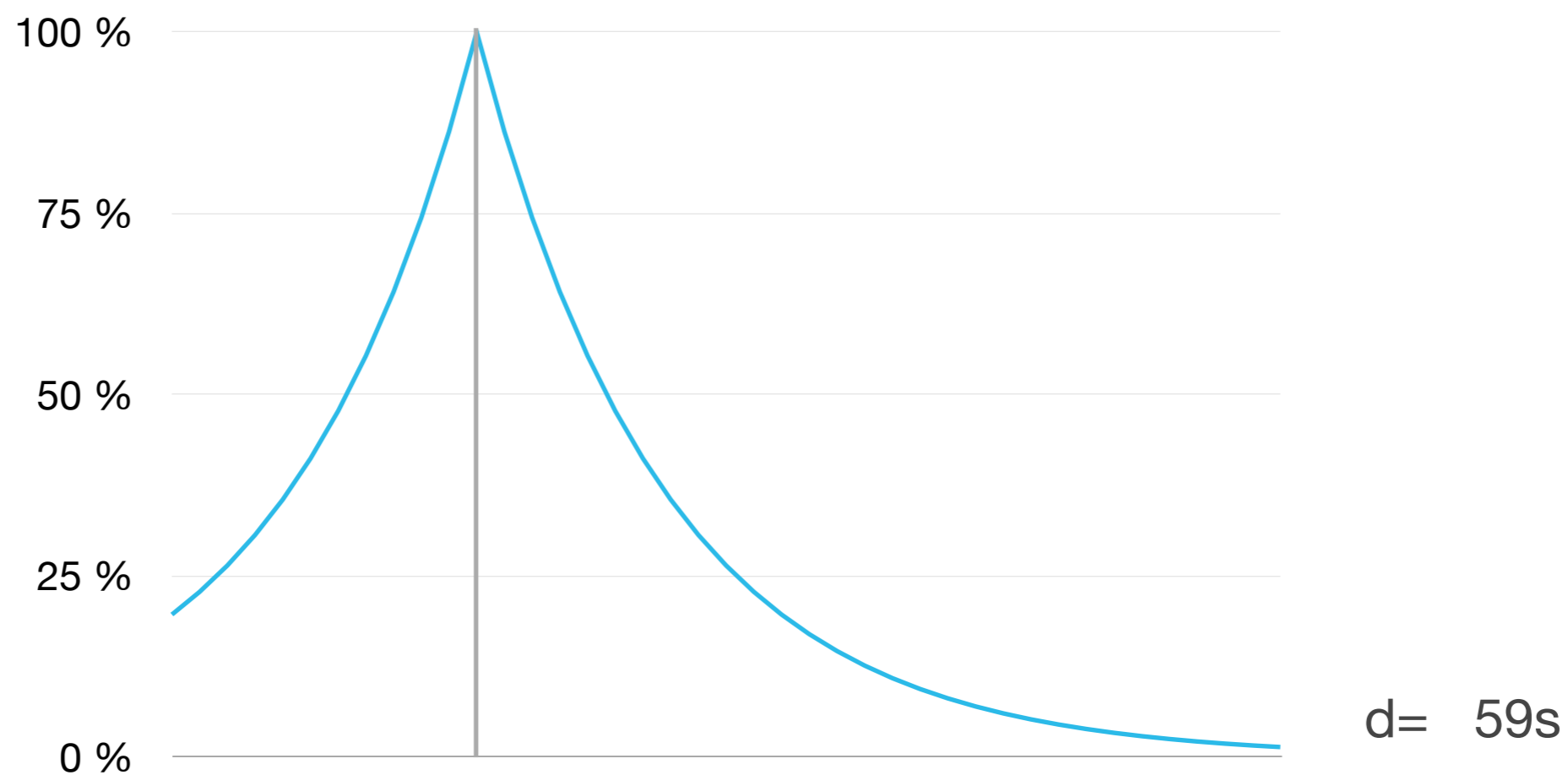
- ▶ Defining visits : threshold ?



drawbacks

Two problems :

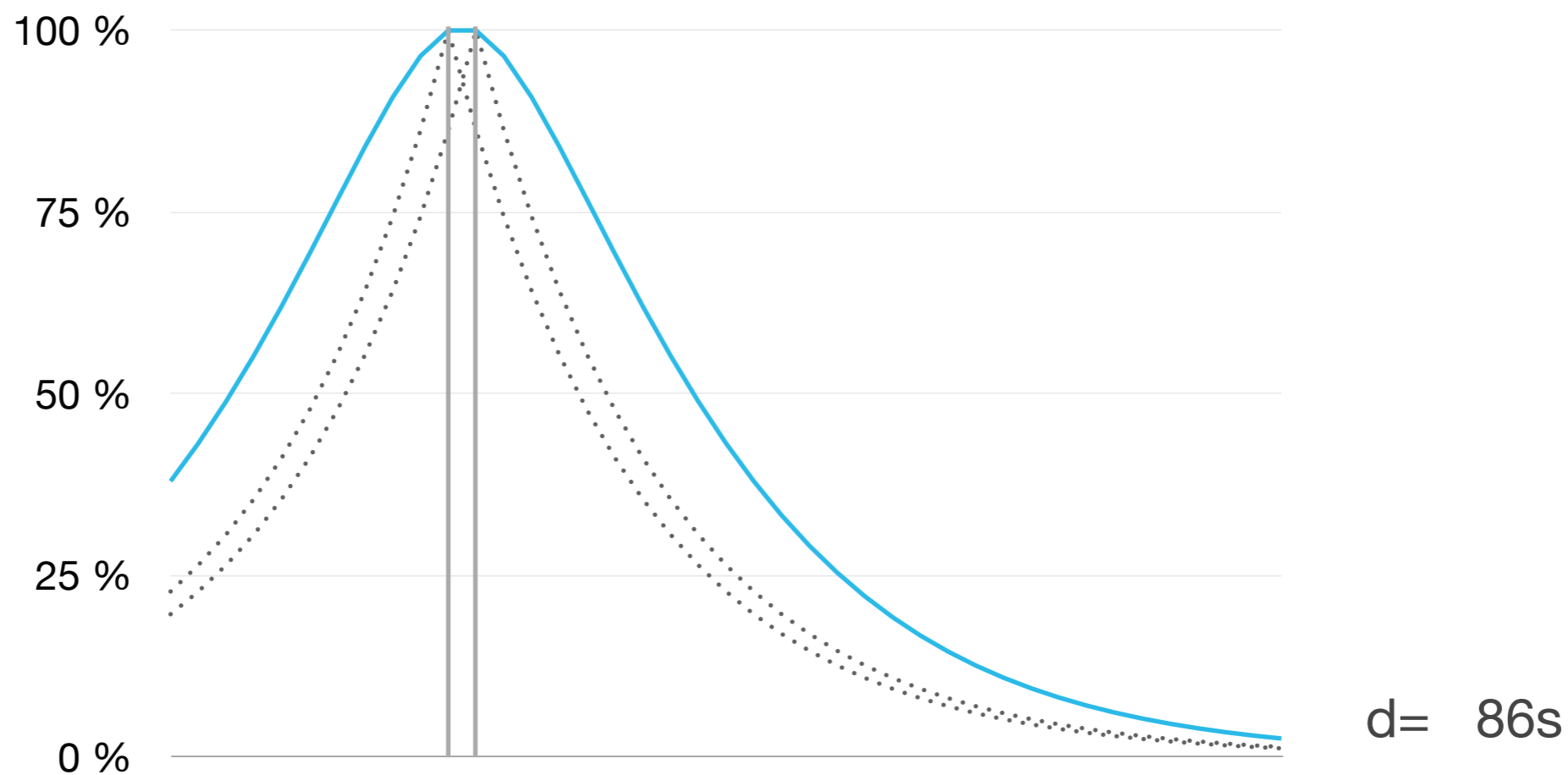
- ▶ Defining visits : threshold ?
- ▶ Overestimating duration.



drawbacks

Two problems :

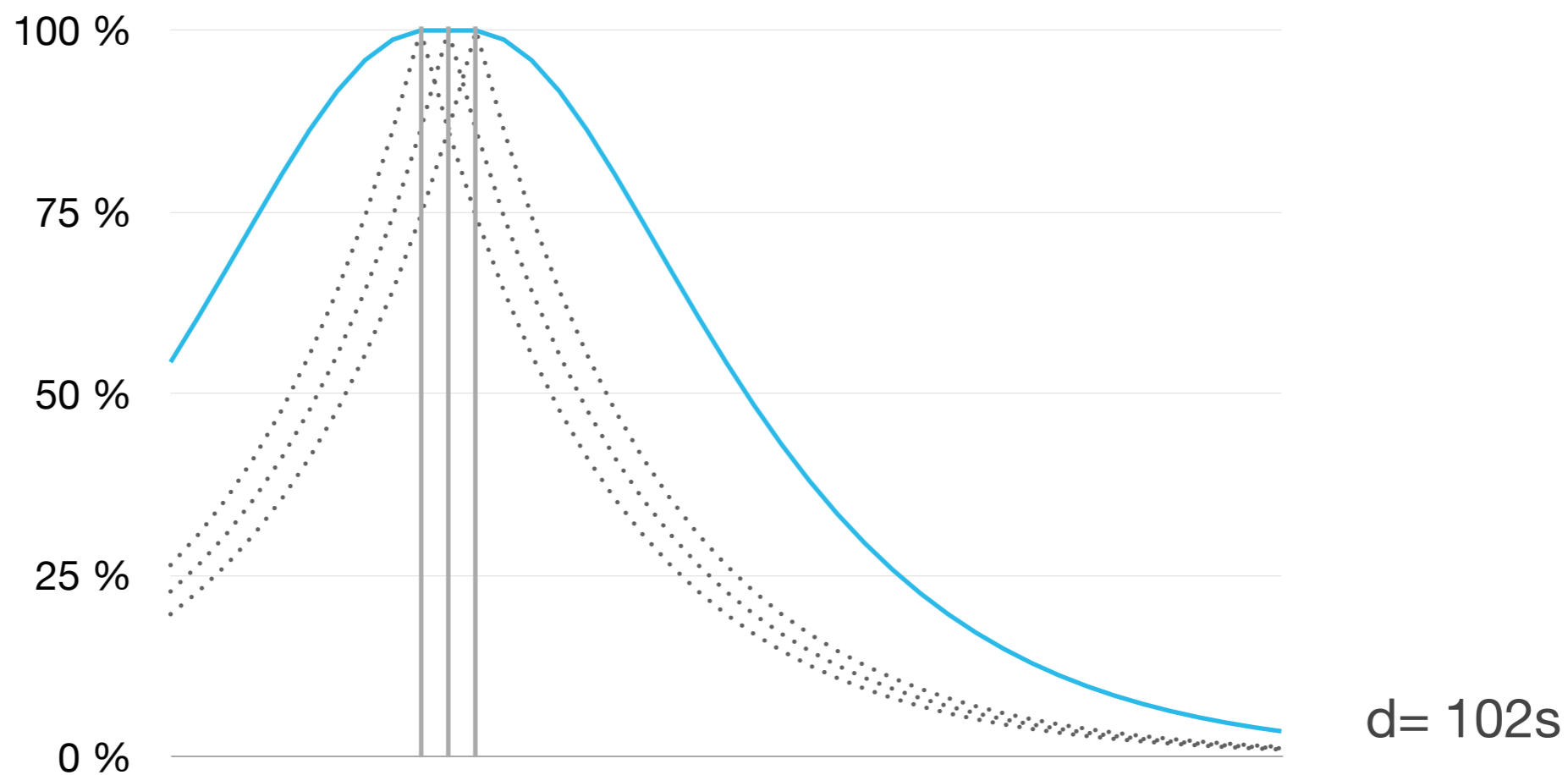
- ▶ Defining visits : threshold ?
- ▶ Overestimating duration.



drawbacks

Two problems :

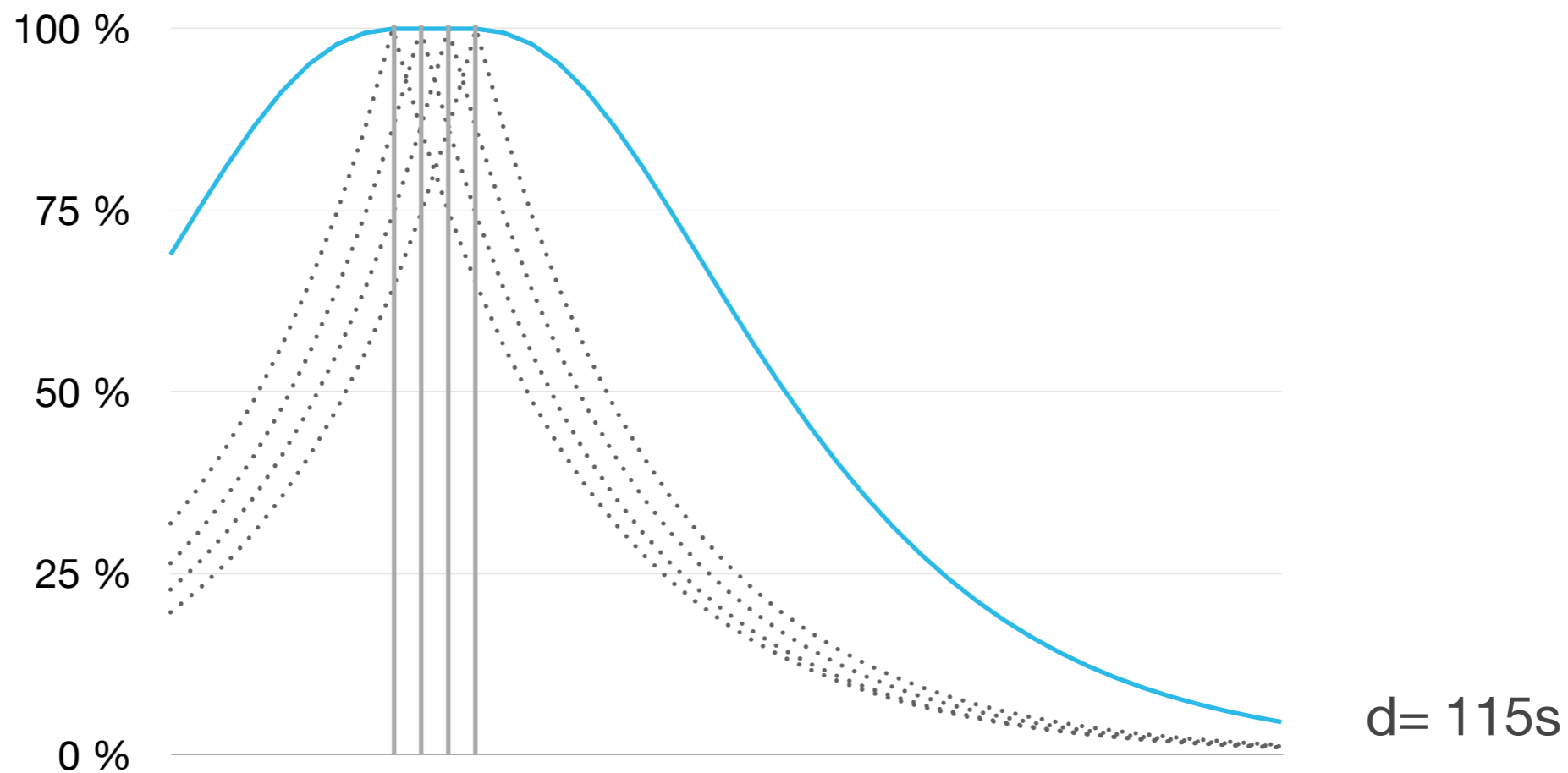
- ▶ Defining visits : threshold ?
- ▶ Overestimating duration.



drawbacks

Two problems :

- ▶ Defining visits : threshold ?
- ▶ Overestimating duration.



gap filling

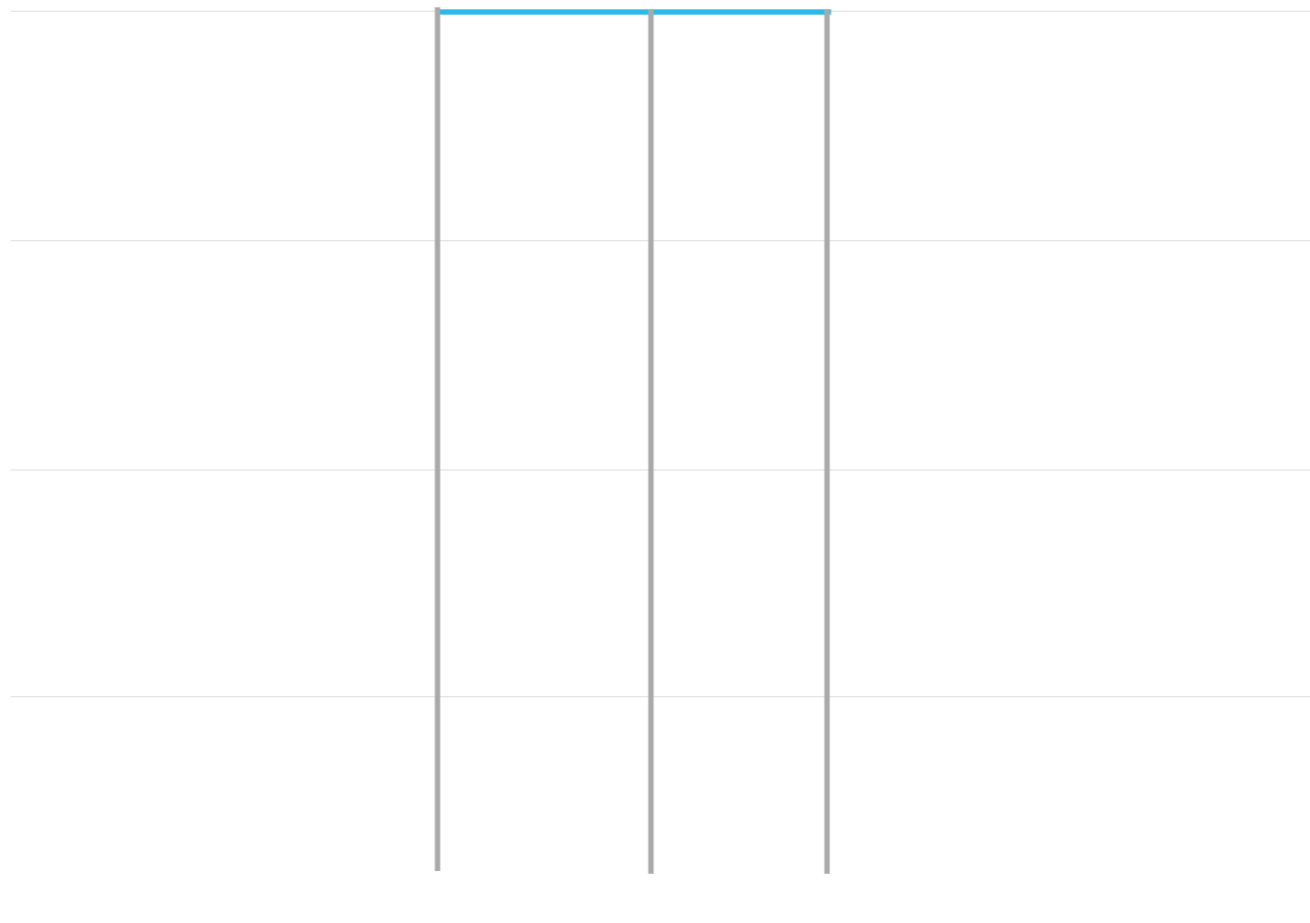
Simpler approach :

- ▶ n contacts

gap filling

Simpler approach :

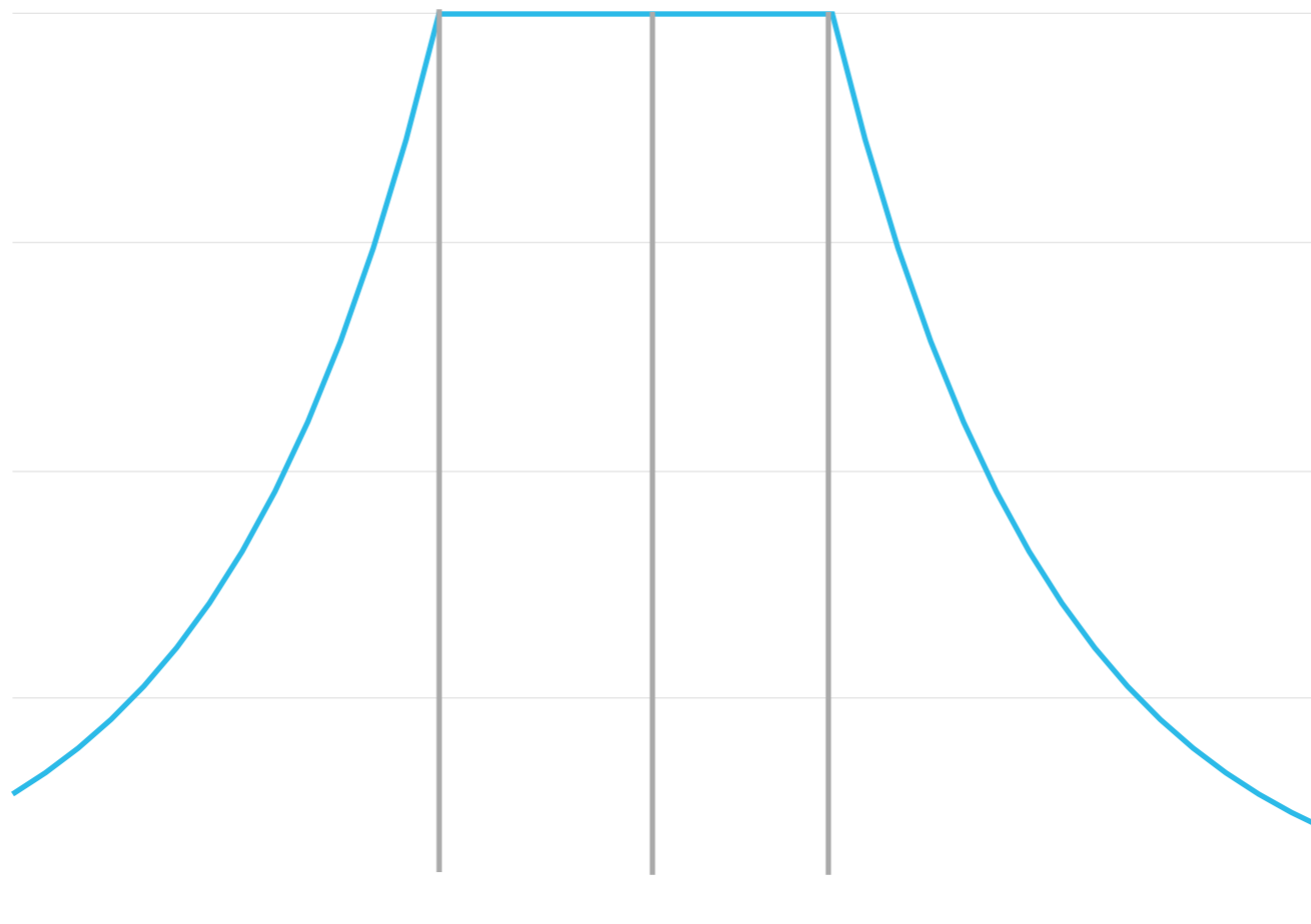
- ▶ n contacts
- ▶ fill *gaps* if intercontact $< d$



gap filling

Simpler approach :

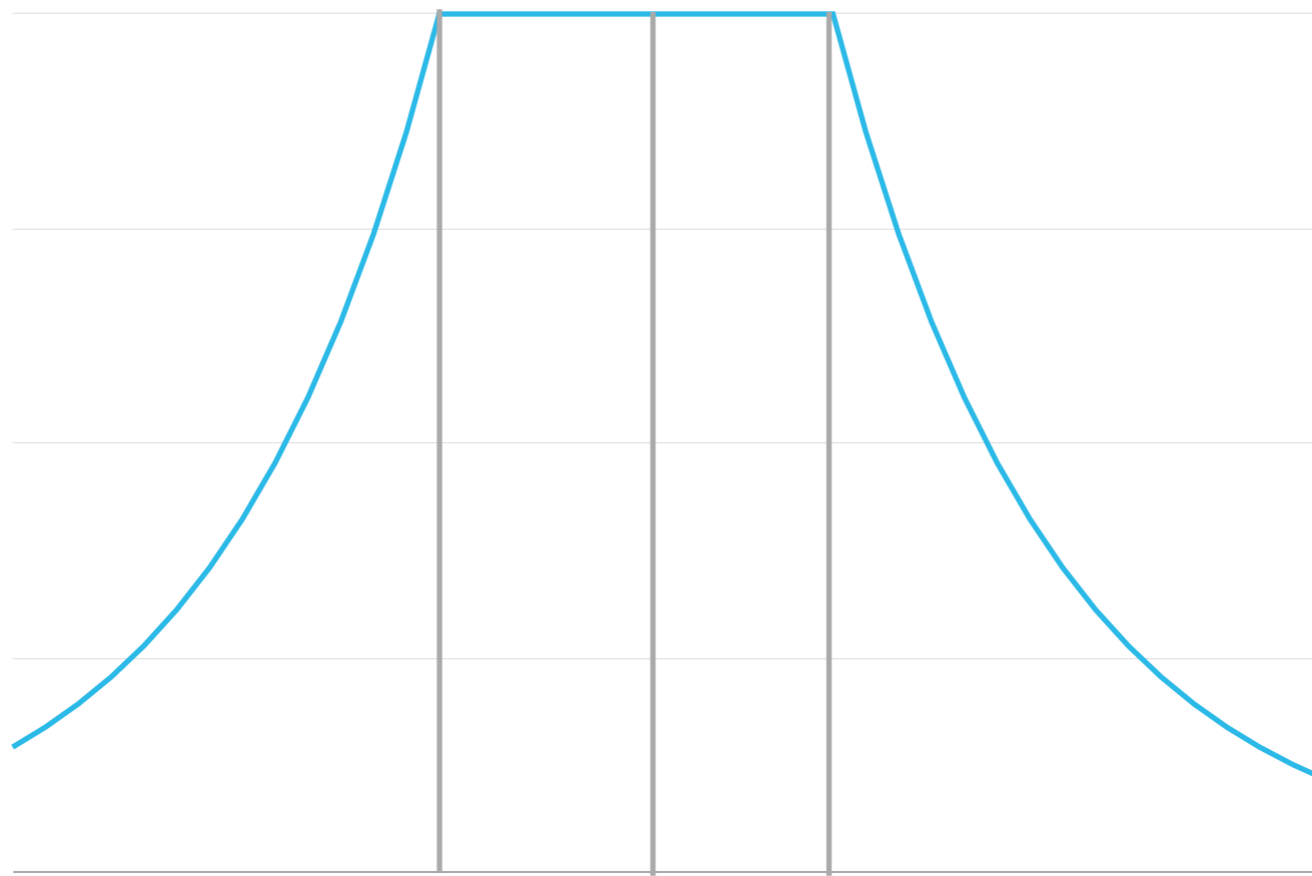
- ▶ n contacts
- ▶ fill *gaps* if intercontact $< d$
- ▶ probabilistic reconstruction on the edges



gap filling

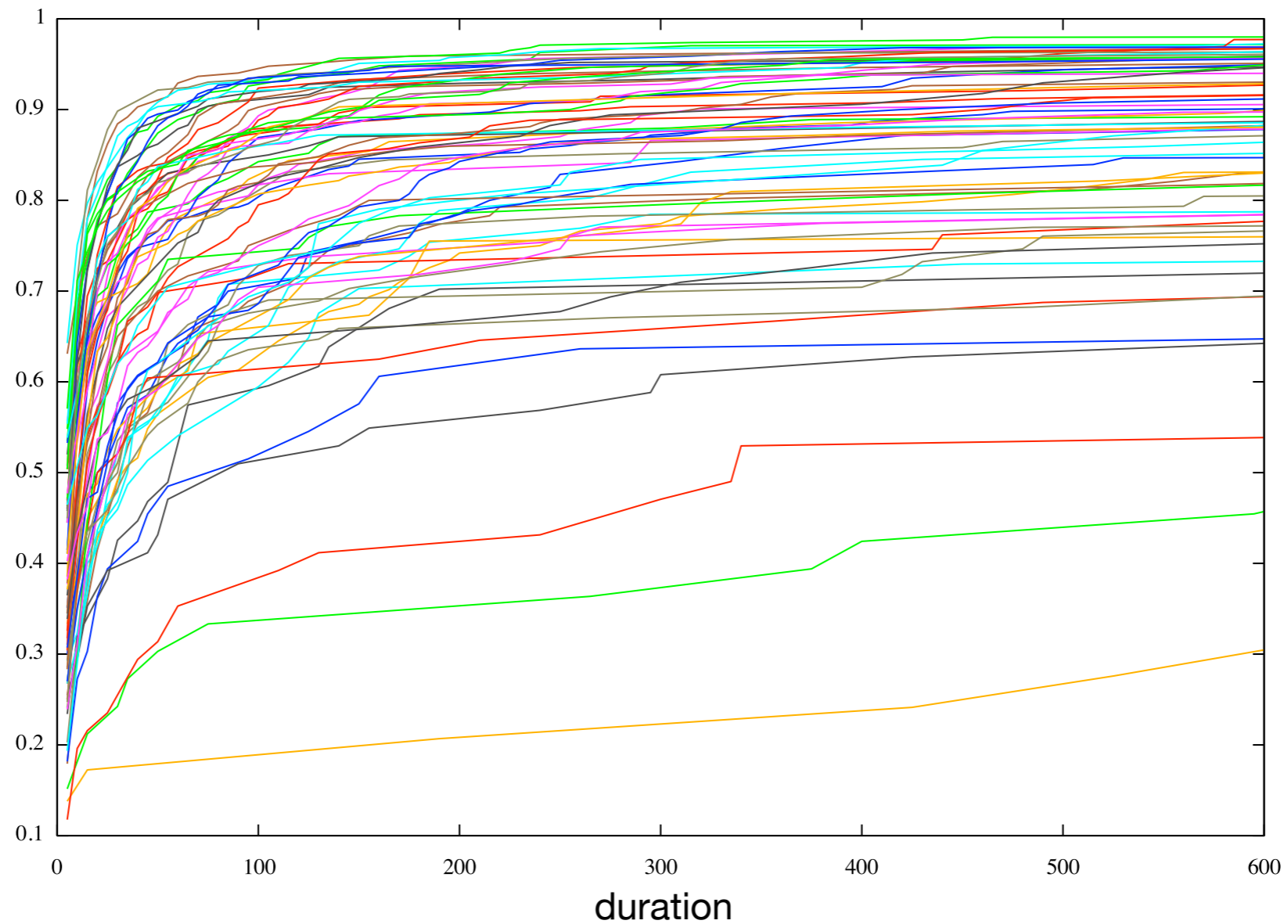
Simpler approach :

- ▶ n contacts
- ▶ fill *gaps* if intercontact $< d$ ← **which d ?**
- ▶ probabilistic reconstruction on the edges



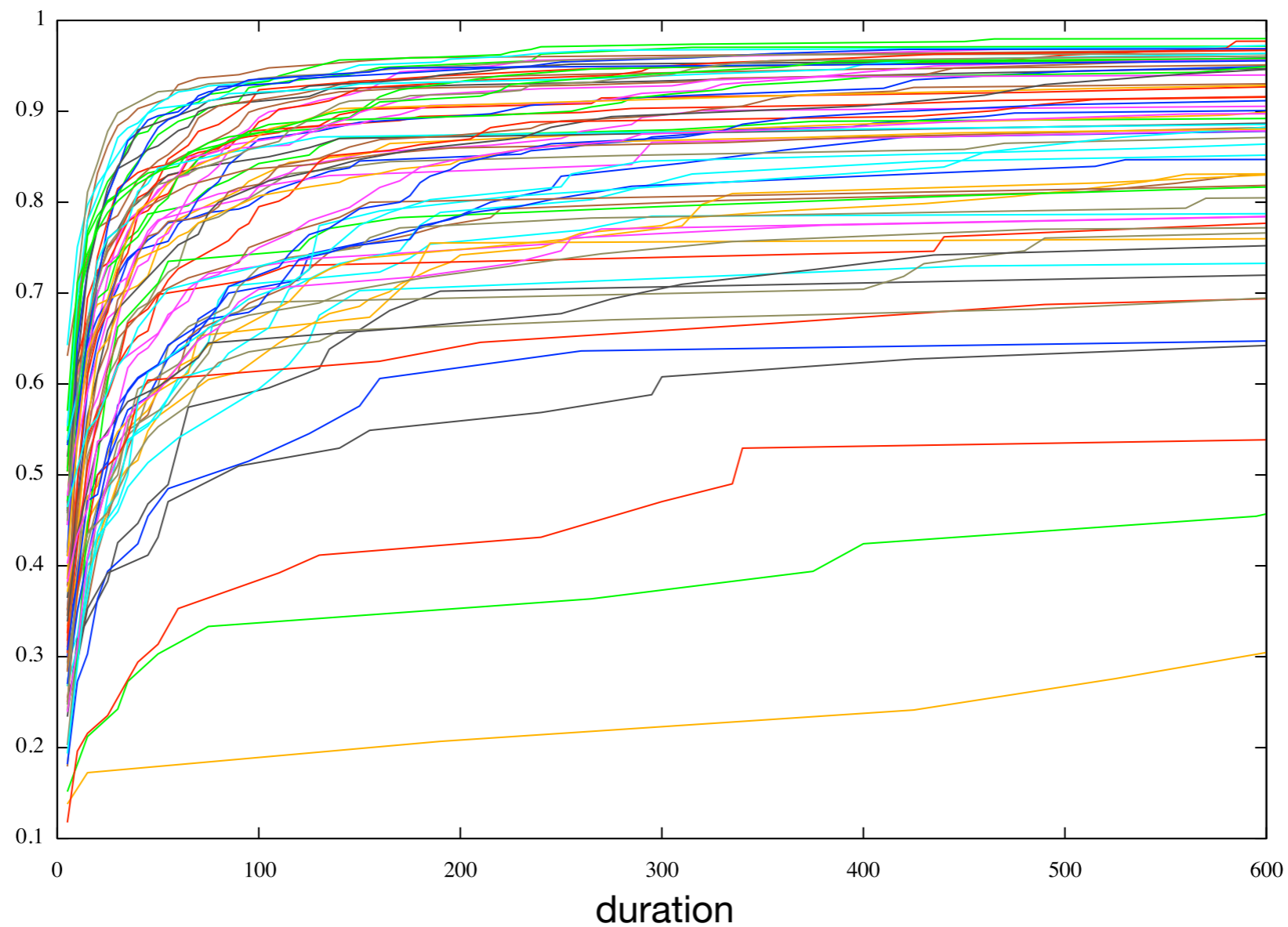
Intercontact duration p.d.f

- ▶ flattens at ~180s

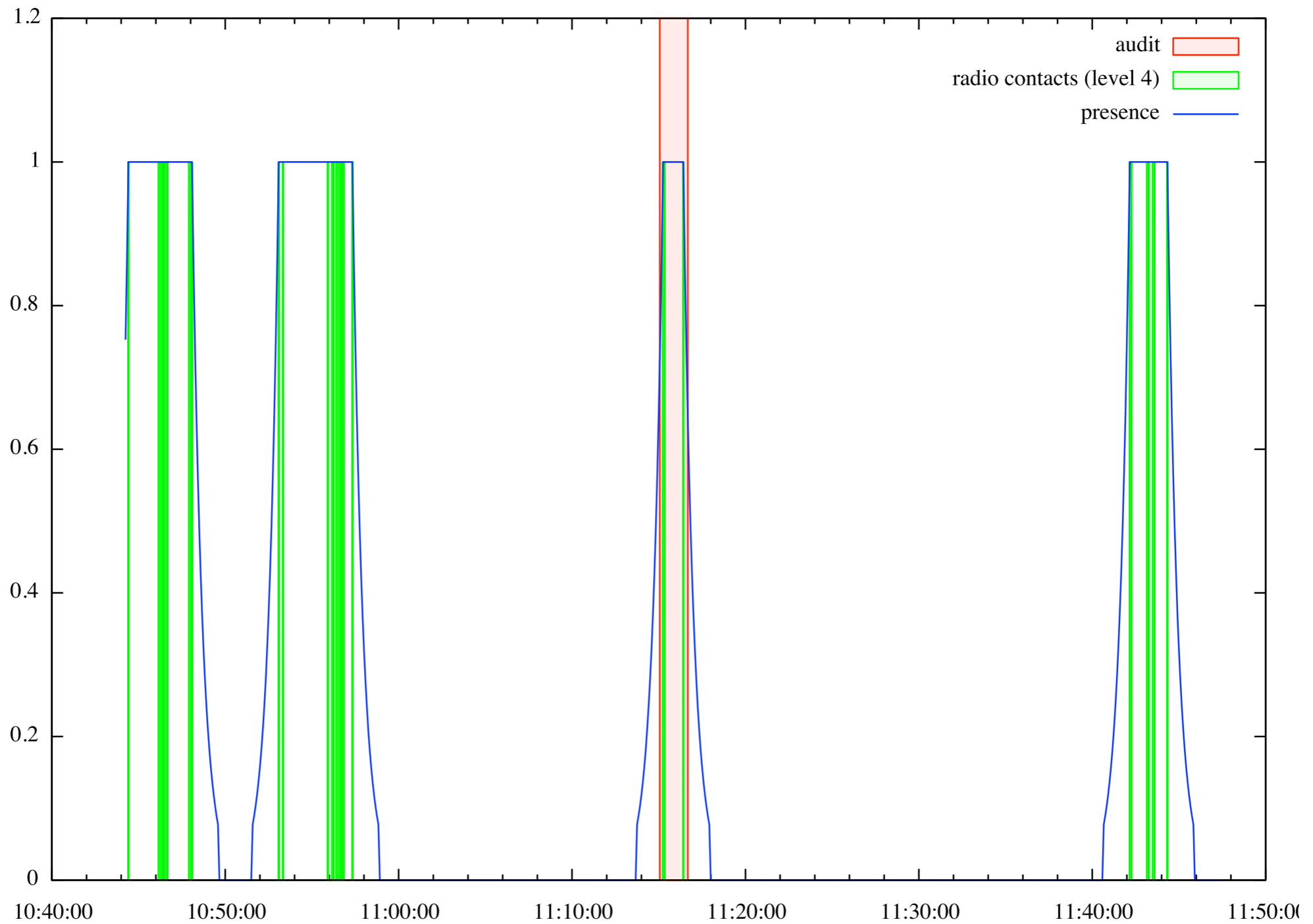


Intercontact duration p.d.f

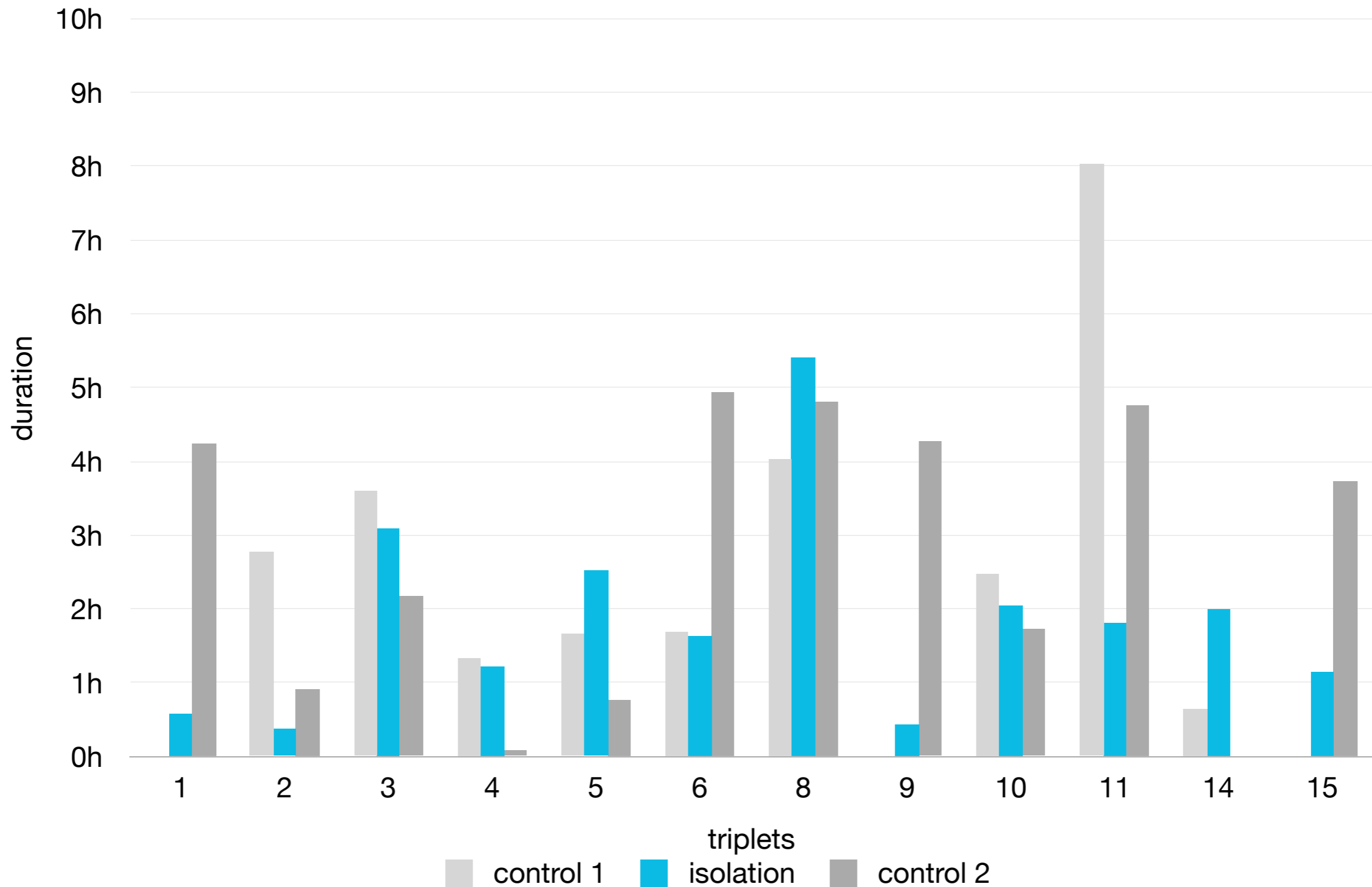
- ▶ flattens at ~ 180 s $\rightarrow d=180$



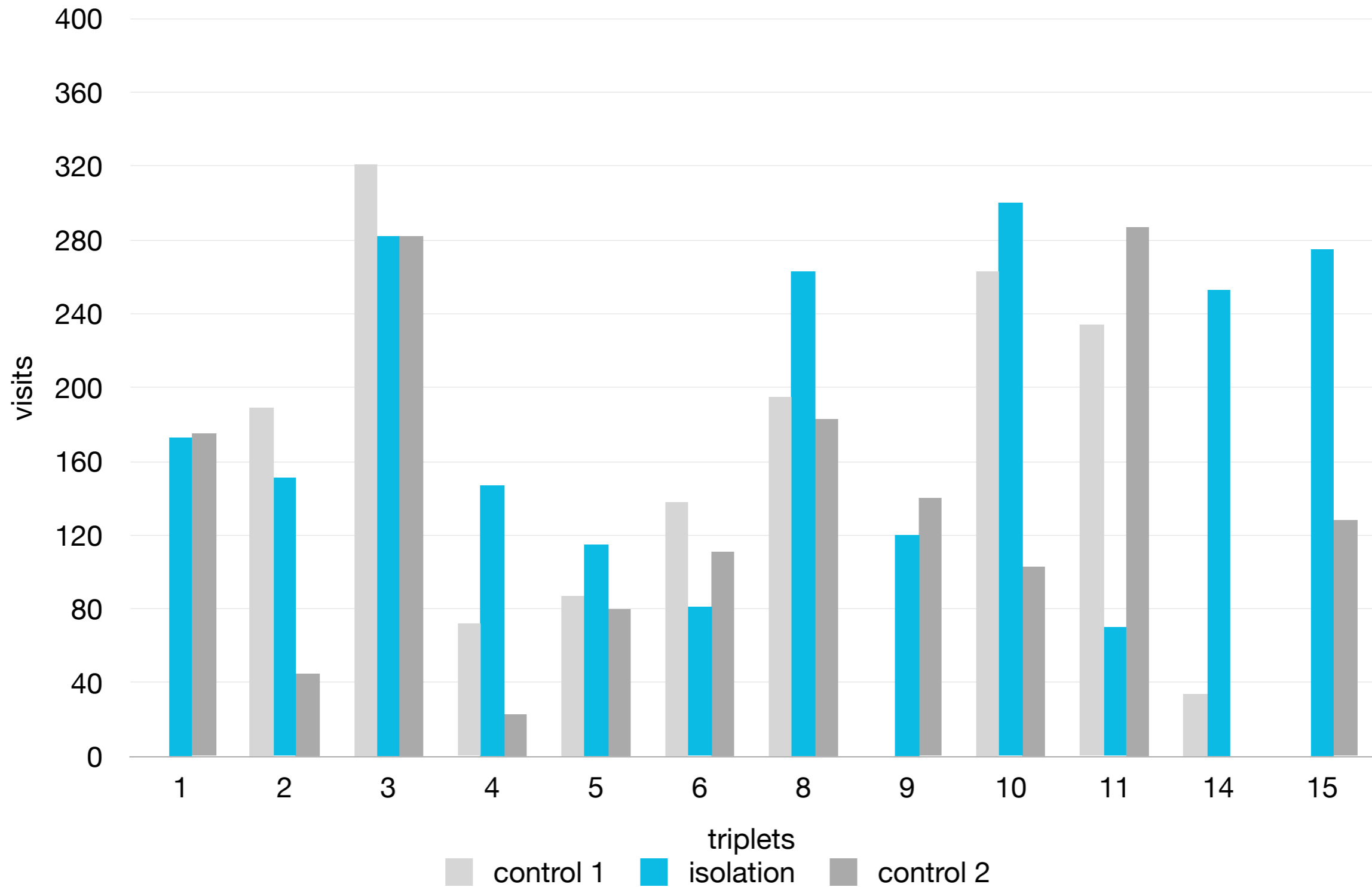
gap filling



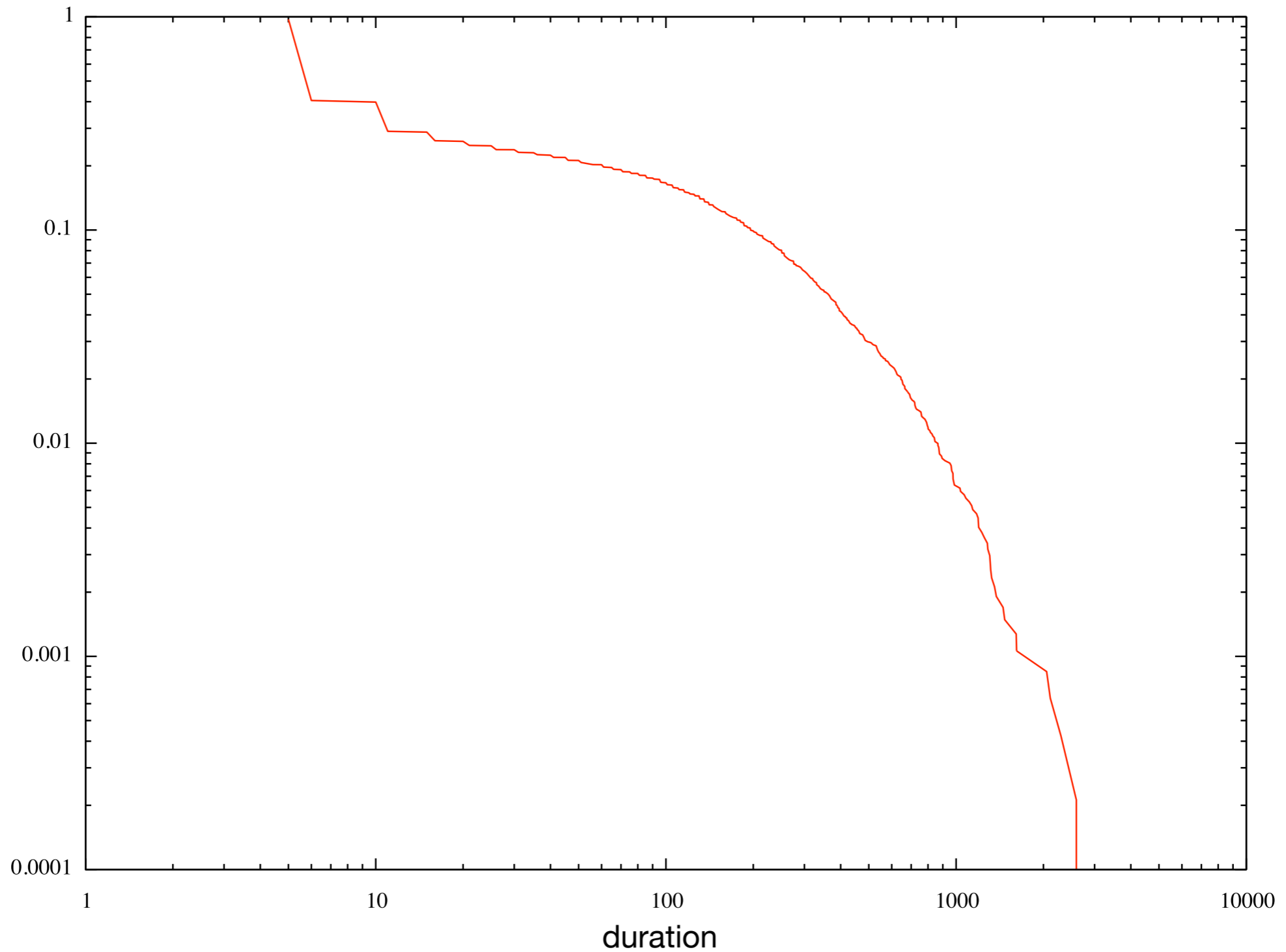
results (durations)



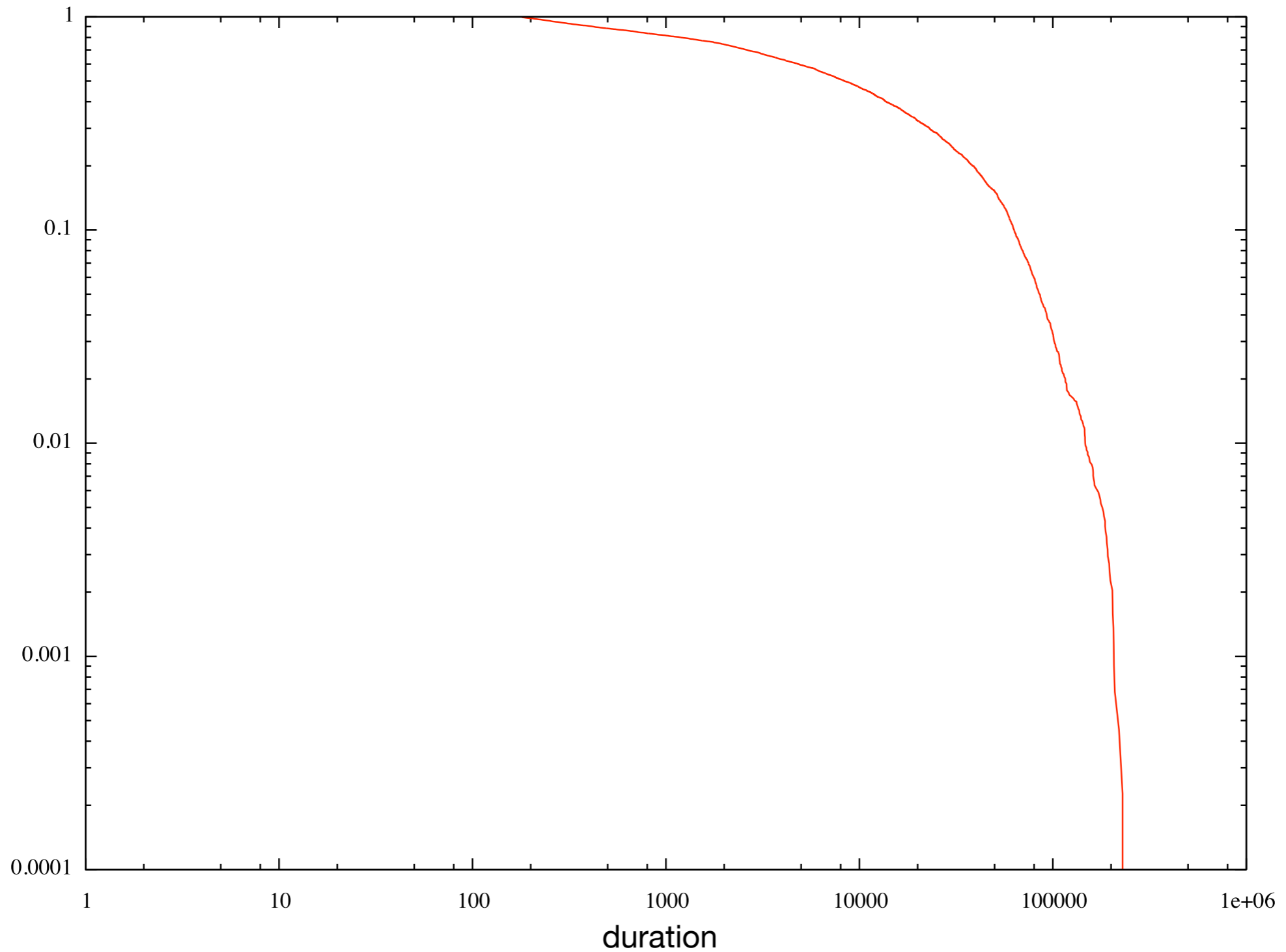
results (visits)



durations c.d.f



intercontacts c.d.f

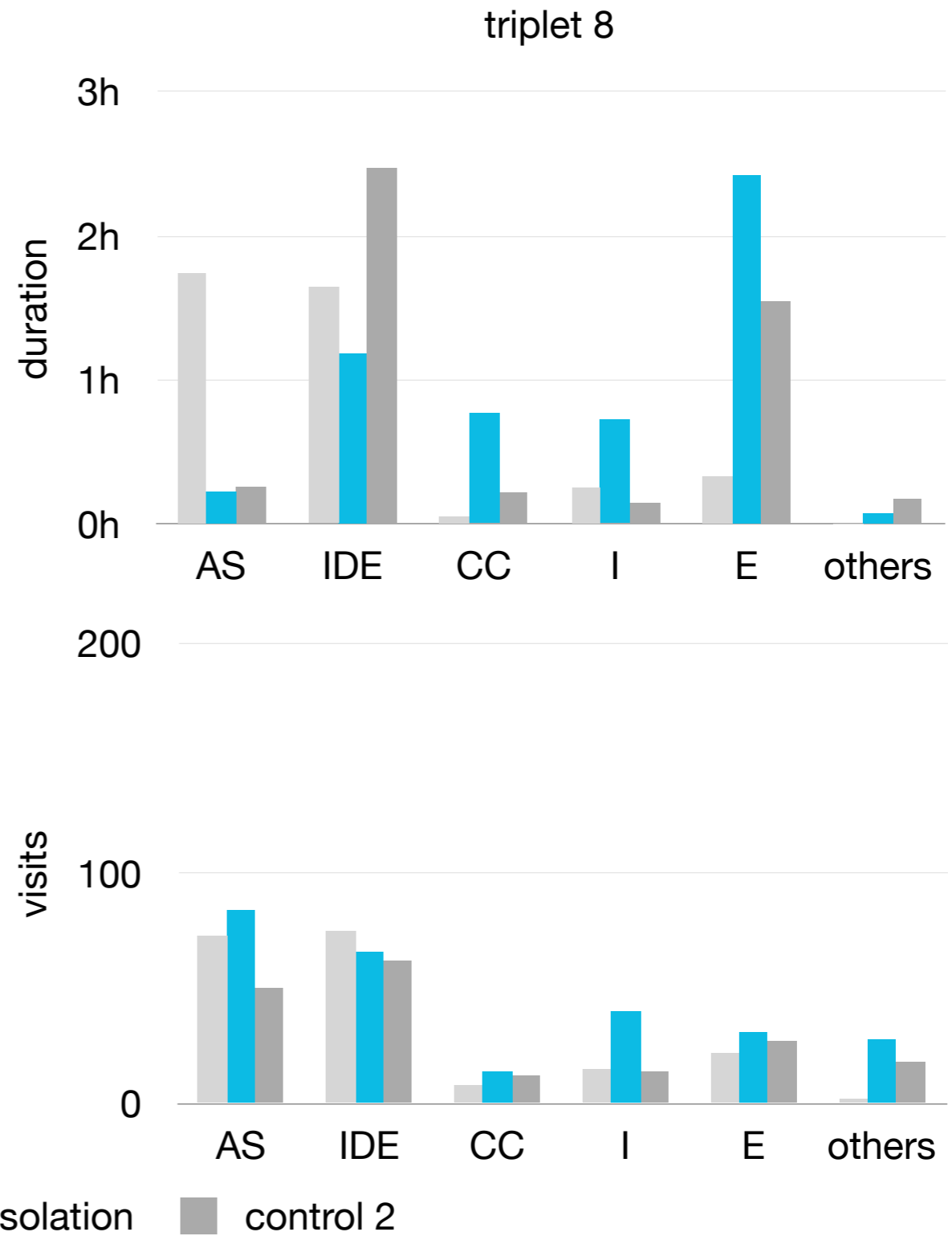
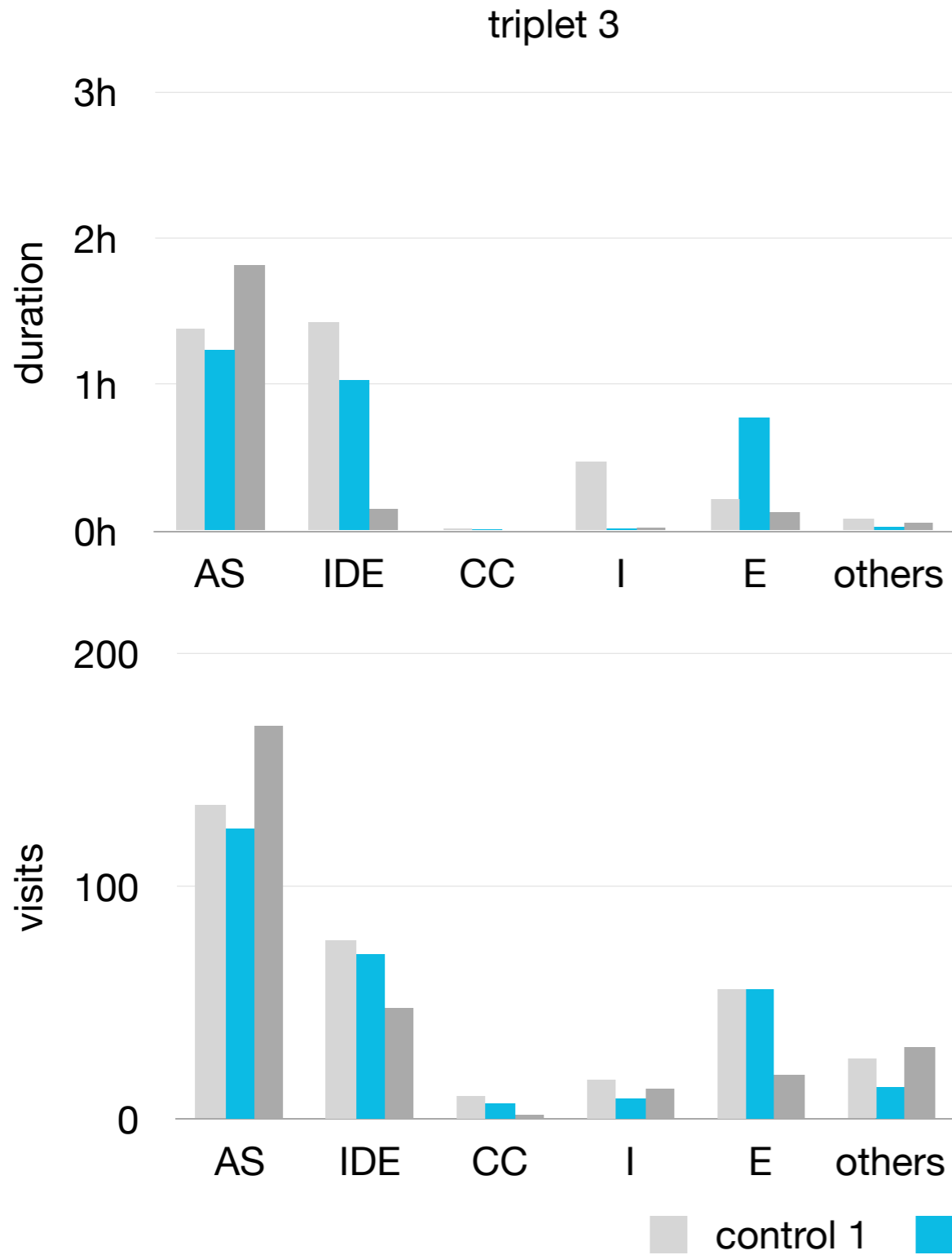


occupations

Finer grain, 6 occupations :

- ▶ AS
- ▶ IDE
- ▶ CC
- ▶ I
- ▶ E
- ▶ others

results (per occupation)



conclusion

- ▶ data corrupted by packet loss
- ▶ several reconstruction methods
 - ▶ simpler is better
- ▶ observations coherent with “expected” results
- ▶ heterogeneous behaviors
 - ▶ next : per occupation behavioral profiles