

Learning network states from RTT measurements

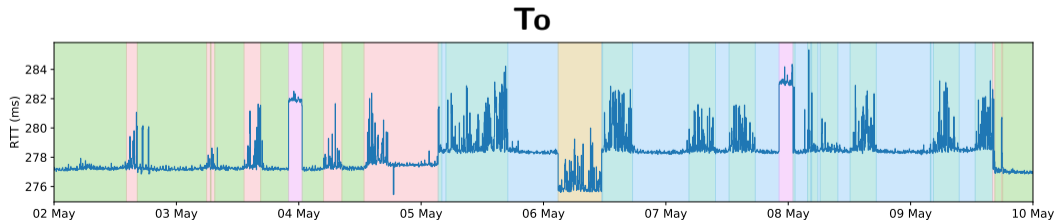
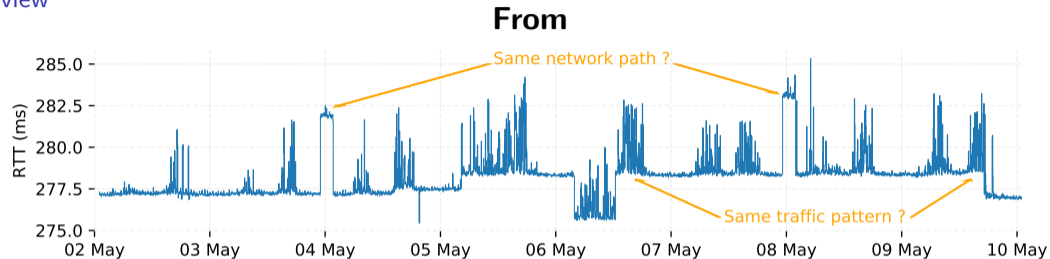
(Briefly)

M. Mouchet T. Chonavel S. Vaton
E. Aben J. den Hertog

IMT Atlantique; RIPE NCC

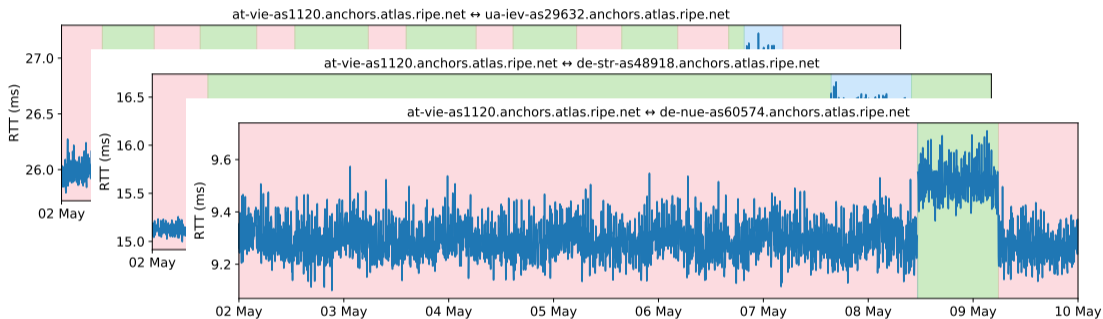
Inferring the Internet's hidden network states

Overview



Inferring the Internet's hidden network states

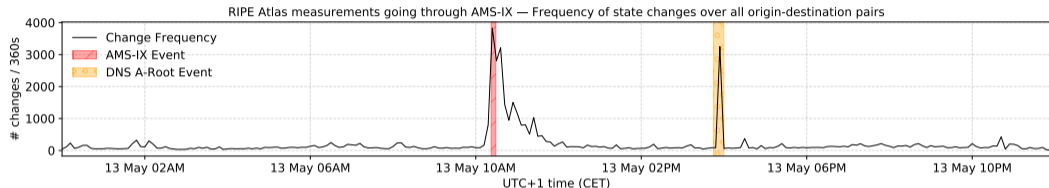
Anomaly detection



1. Learn the model for *selected* origin-destination pairs
2. Extract change points (state changes)
3. Compute the change frequency

Inferring the Internet's hidden network states

Anomaly detection



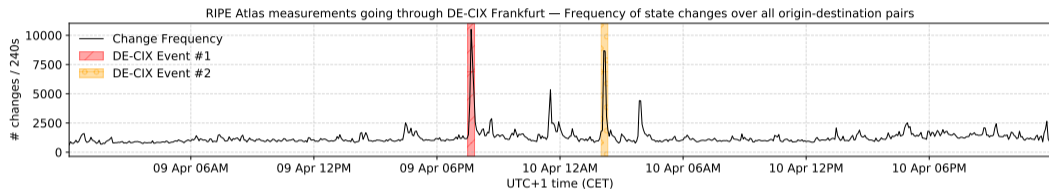
Change frequency on the 13th of May 2015 for the 20k pairs that saw AMS-IX Frankfurt in their traceroutes the day before.

”[...] the primary cause of **the event lasted for seven minutes and two seconds**, from 2015-05-13T12:22:12+02:00 to 2015-05-13T12:29:14+02:00. This was the interval from a switch interface starting to generate **looped traffic** until the backbone interface of the switch was shut down [...]”¹

¹<https://labs.ripe.net/Members/emileaben/does-the-internet-route-around-damage>

Inferring the Internet's hidden network states

Anomaly detection



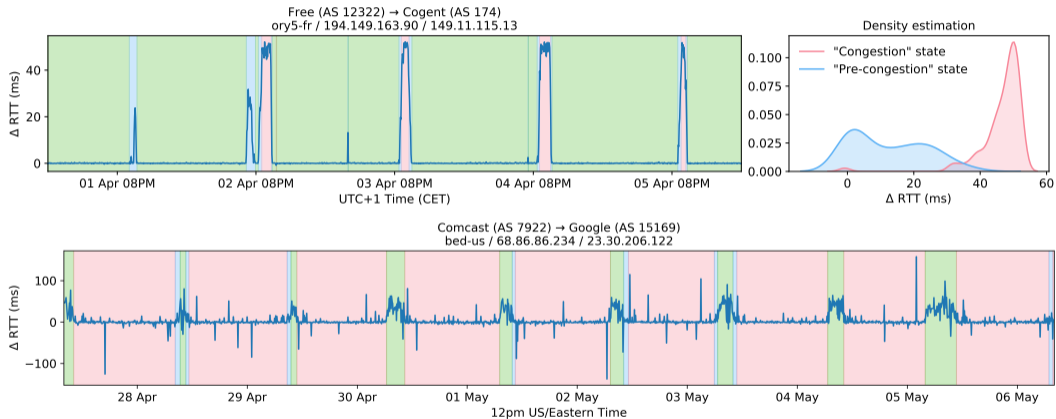
Change frequency between the 9th and the 10th of April 2018 for the 60k pairs that saw DE-CIX Frankfurt in their traceroutes the day before.

”[...] there were **2 separate outage events**. One from 19:30-23:30 on 9 April, the other from 02:00-04:00 on 10 April (all times UTC). [...] From the information we gathered, we think these are the times that many networks **lost connectivity to the DE-CIX route-servers** [...]”²

²<https://labs.ripe.net/Members/emileaben/does-the-internet-route-around-damage-in-2018>

Inferring the Internet's hidden network states

Congestion detection with MANIC data



Segmentation of RTT difference (far - near) time series from the CAIDA MANIC project (one color per state).

Inferring the Internet's hidden network states

Reference implementation

```
using RTTHMM
```

```
timestamps, observations = read_measurement("1001_1_193.0.14.129.csv")  
# => [1548806560, 1548806786, 1548807026, 1548807273]  
# => [49.53, 50.75, 48.79, 49.58, ...]
```

```
stateseq = fit(observations, timestamps)  
# [Thread 1] mean(dt) = 240.0, std(dt) = 0.0  
# [Thread 1] Clustering 466 observations...  
# => [2,2,2,...,1,1]
```

(Not yet public :-)

<https://github.com/maxmouchet/RTTHMM.jl>

Inferring the Internet's hidden network states

RIPE Atlas implementation

```
{
  "n_states": 5,
  "states": {
    "1": {
      "rtt": {
        "max": 224.688,
        "median": 206.499,
        "iqr": 0.133,
        "min": 199.882
      },
      "duration": {
        "total_time": 145680,
        "avg_time": 72840.0
      }
    }, "..."
  },
  "segments": [
    {
      "start": 0,
      "stop": 64,
      "start_time": 1550448206,
      "stop_time": 1550463568,
      "state": "1"
    },
    {
      "start": 64,
      "stop": 408,
      "start_time": 1550463568,
      "stop_time": 1550546126,
      "state": "3"
    }, "..."
  ]
}
```

<https://github.com/maxmouchet/atlas-trends-demo>

Inferring the Internet's hidden network states

References

- ▶ M. Mouchet, T. Chonavel, and S. Vaton. **Statistical Characterization of Round-Trip Times with Nonparametric Hidden Markov Models**. In *IFIP/IEEE IM 2019 Workshop: 4th International Workshop on Analytics for Network and Service Management (AnNet 2019)*. Washington DC, USA.
- ▶ M. Mouchet, S. Vaton, T. Chonavel, E. Aben, and J. den Hertog. **Large-Scale Characterization and Segmentation of Internet Path Delays with Infinite HMMs**. *Preprint* (Oct. 2019).

Inferring the Internet's hidden network states

Challenges

- ▶ Interpretation of the hidden states ? Congestion, route changes, ...
- ▶ Inference (of the hidden states) scales linearly with the number of observations, but there's an high constant cost.
 - ▶ $\sim 3s$ for 1 week of mesh measurements (2520 RTT observations) on an recent AVX2-enabled CPU.
 - ▶ Online learning as a solution (do not re-learn everything as we add observations).
- ▶ Anomaly detection at scale ?
 - ▶ Too many pairs ($523^2 \times 3s \simeq 900000s \simeq 10\text{days}$), how to choose where to look ?
 - ▶ How to find (easily) Atlas pairs going through a specific AS/IXP ?