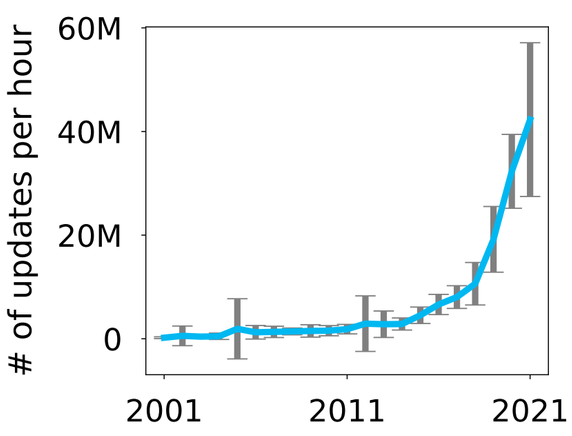
MVP : Measuring Internet Routing from the Most Valuable Points

Thomas Alfroy, *Thomas Holterbach, Cristel Pelsser*

Background

- Internet routing is monitored by BGP Vantage Points (VPs) [1,2]
- RIPE RIS infrastructure includes 1000 VPs, that collect more than **150GB per day**

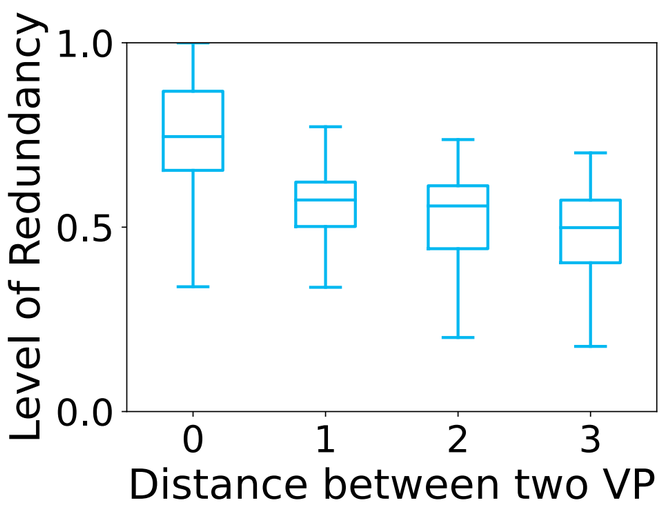
Problem



Exponential increase of the volume of data

Redundancy also increases because the location of VPs is skewed [3]

Location Similarity ⇏ Data Redundancy



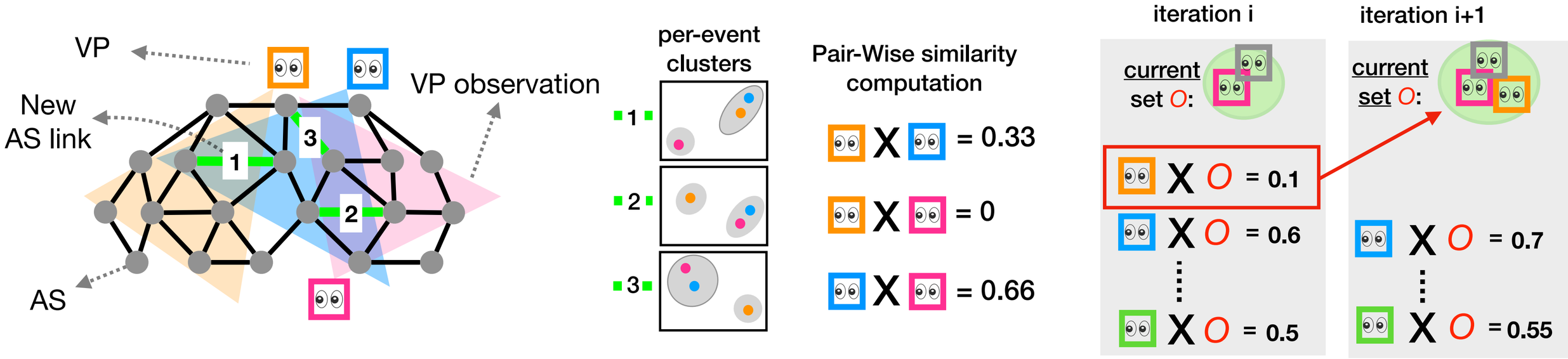
High distance does not systematically mean low redundancy

AS distance is not a good metric to estimate data redundancy for two reasons:

- **Incomplete** inferred AS topology
- **Hidden** Routing Policies

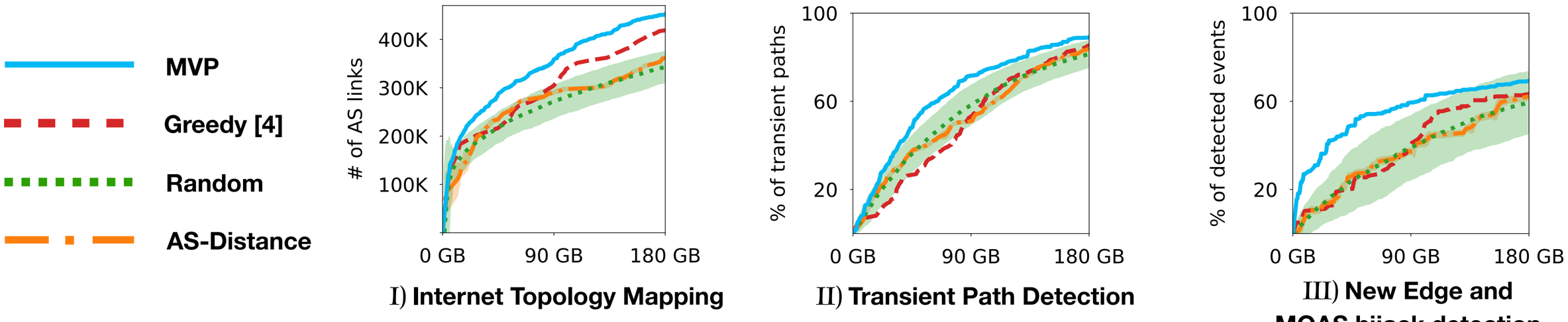
MVP finds VPs that provides dissimilar data, based on historical information

1. We use past BGP events to quantify VPs observation, using topological features
2. We compute a *Pair-Wise Similarity Score* using a clustering algorithm
3. We *greedily build* a set of dissimilar VPs



MVP strikes the best balance between utility and volume of data

We evaluate MVP on three possible use cases and compare it against three baselines.
We focus on the tradeoff between *utility* and *volume of data.*



MVP returns the best tradeoff between *Utility* and *Volume* for every use case.

Acknowledgement

Thomas Alfroy is funded by ArtIC project "Artificial Intelligence for Care" (grant ANR-20-THIA-0006-01) and co funded by Région Grand Est, Inria Nancy - Grand Est, IHU of Strasbourg, University of Strasbourg and University of Haute-Alsace.

Thomas Holterbach is partially funded by ISOC, the MANRS Observatory and by Région Grand Est.

References

[1] 2016. RIPE RIS Raw Data. https://www.ripe.net/data-tools/stats/ris/. (2016).

[2] University of Oregon. 2016. Route Views Project. (2016). www. routeviews.org/.

[3] Bias in Internet Measurement Infrastructure, https://ripe84.ripe.net/archives/video/768/

[4] Ying Zhang, Zheng Zhang, Z. Morley Mao, Y. Charlie Hu, and Bruce M. Maggs. 2007. On the impact of route monitor selection. In IMC '07.