

Evaluation of Network Categorization Strategies

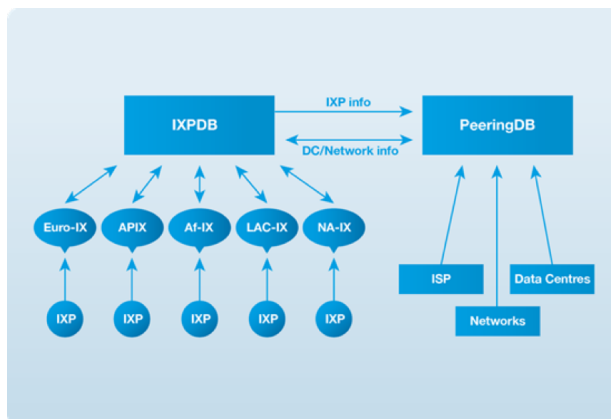
Motivation

Internet measurements often include scanning through enormous amounts of targets, without having the ability to understand the exact characteristics or function of each scan target. The scanner targets Internet routers, content servers, home router boxes and Internet routers alike while the researcher do not have the resources to analyze this difference for each individual target.

An important step to better understand scan results on a relatively high level without too much effort are network category databases.

Existing databases include the community-maintained PeeringDB [1], ASdb [2], IP2Location [3] and others [4].

The main aim of this thesis is to develop and evaluate a mechanism to categorize our own scanning data to help us better understand it.



Your Task

- Familiarize yourself with the existing data sources
- Apply them to our own scan data [5] and develop metrics to measure the categorization accuracy
- Develop your own, combined categorization strategy from your insights and apply it to our scan data

References

- [1] <https://www.peeringdb.com/>
- [2] <https://zakird.com/papers/asdb.pdf>
- [3] <https://www.ip2location.com/database/ip2location>
- [4] <https://prefixtoplists.net.in.tum.de>
- [5] <https://ipv6hitlist.github.io>

Requirements

Basic knowledge or willingness to learn about internet infrastructure and routing mechanisms. Terms like ASN, CDN or BGP should be well known.

Contact

Lion Steger stegerl@net.in.tum.de
Johannes Zirngibl zirngibl@net.in.tum.de
Patrick Sattler sattler@net.in.tum.de

