Technische Universität München, Department of Informatics

Chair for Network Architectures and Services
Prof. Dr.-Ing. Georg Carle

Thesis B.Sc.

Thesis M.Sc.

Parsing geographical locations from DNS names - at scale!

Motivation

Internet measurement studies frequently come along DNS names that include city names or airport codes. Frequently, these are being used to assign a geographical location to a node, though this approach has been proven partially unreliable by Zhang et al. [1]. This work aims to extend the existing work by having a larger scale and an evaluation of accuracy of the geographical identifiers found.

```
ae1.br01.fra1.tfbnw.net
be2.bb01.fra2.tfbnw.net
be2.bb02.fra2.tfbnw.net
ae12.bb01.ams2.tfbnw.net
ae2.bb01.ams2.tfbnw.net
ae3.bb02.bos2.tfbnw.net
ae9.bb02.lhr2.tfbnw.net
be9.bb01.ewr2.tfbnw.net
```

Frankfurt, Amsterdam, Boston, London and New York city codes

Approach

In a first step, reverse DNS resolutions of the entire IPv4 and parts of the IPv6 address space are to be gathered and potentially expanded by own measurements.

In a second step, these DNS names are then to be scanned for geographical codes embedded into these names (e.g. FRA for Frankfurt).

In a third step, the geographical location of those nodes is to be compared against existing geolocation approaches (e.g. [2]) and against own measurements through e.g. PlanetLab or Ripe Atlas to verify geographical location.

As a stretch goal, a fourth step could include correlation of that data to other opportunistic sources of geolocation, e.g. whois data, TLS certificate data or BGP community strings.

[1] M. Zhang et al., How DNS Misnaming Distorts Internet Topology Mapping, *USENIX Conference 2006*

[2] MaxMind, GeoLite City databases, http://dev.maxmind.com/geoip/geoip2/geolite2/

Contact

Quirin Scheitle scheitle@net.in.tum.de
Oliver Gasser gasser@net.in.tum.de
Johannes Naab naab@net.in.tum.de

http://go.tum.de/644204