# CS 8803 Internet Data Science

Alberto Dainotti 1/13/2022



# Welcome to Day 10



# Internet protocols and architectures

- The TCP/IP stack, application and network protocols, and (Inter)networking architectures are important
- They make us understand how packets travel and how applications can communicate, etc.



https://sdn.systemsapproach.org/arch.html

# **Internet Data Science?**

But if we want to empirically understand how human behavior and natural phenomena affect or inter-relate with the Internet—or if we want to study the Internet itself—then we need to add something else to the picture..

.. and use a different perspective

To help us reason about this different perspective we can get back to the drawing board...



# **Internet Data Science?**

- Each Internet-related phenomenon can be observed along some of these dimensions
  - IP addresses (individual or aggregated per subnet, etc.)
  - Protocol types and port numbers
  - Domain names
  - Autonomous System Numbers
  - Organizations involved
  - Infrastructure used
  - Geopolitical dimensions (coordinates, or administrative boundaries: countries, cities, ...)
  - Time

































Fig. 1: Maps of the 204 metro fiber networks in the continental US used in our study. Magnified areas shown include Portland, Chicago, Phoenix, and Miami. *"Characteristics of Metro Fiber Deployments in the US", P. Barford et al., TMA 2020*<sup>20</sup>





















# **Internet Data Science**

We know these entities exist, but

- Do we have data about them? What kind?
- Do we understand the relations among them?
- What data is available about these relations?
- Do we understand these data?
  - Quality and limitations
  - Comprehensiveness
  - Exceptions
  - Validity over time, ...
  - Can we obtain/extract more?
- What can we **do** with it?











# **These relationships**

- Complex: not 1-1, not defined, heterogenous behavior, ...
- Not always well understood
- Scalability issues
- Change over time
- Scarcity of data
- Unofficial data / or not designed for our purposes
- Quality of the data

• In some cases even the entity in those bubbles are not entirely known





# **IP Geolocation: when things get ugly**





# IP Geolocation: when things get ugly



Georgia Tech

https://www.washingtonpost.com/news/morning-mix/wp-content/uploads/sites/21/2016/08/toilet.jpg

# **IP Geolocation: a quiet farmhouse in Kansas...**

### Over several years

- Accused of being identity thieves, spammers, scammers and fraudsters
- Visited by FBI agents, federal marshals, IRS collectors, ambulances searching for suicidal veterans, and police officers searching for runaway children
- The renters have been doxxed, their names and addresses posted on the internet by vigilantes
- Once, someone left a broken toilet in the driveway





https://web.archive.org/web/20170602090701/https://fusion.kinja.com/ how-an-internet-mapping-glitch-turned-a-random-kansas-f-1793856052

# When IP Geolocation means "just the country"



39°50'N 98°35'W 39.8333333,-98.585522

38.0000,-97.0000



https://web.archive.org/web/20170602090701/https://fusion.kinja.com/ how-an-internet-mapping-glitch-turned-a-random-kansas-f-1793856052



















### Select a time range: 😮

3 days ago - Now

### Recent History **v**

### Country Outages ??



Country	11	Overall Score	Active Probing	BGP ↓	Darknet 🗍
Gambia		251M	9.34k	28.3k	
Myanmar		10.4M	9.43k	1.13k	
Equatorial Guinea		469k	909	1.19k	
Europe (Unknown Country)		62.1k		62.1k	
Niger		29.1k	29.1k		
French Guiana		24.9k		24.9k	
Panama		23.3k		23.3k	
Cote D Ivoire		12.7k	12.7k		
Marshall Islands		12.4k	12.4k		
China		11.7k		11.7k	

Showing 1 to 10 of 30 entries

Previous Next



# Scope: large connectivity outages

- Natural disasters
  Cyberattacks
  HW faults
  Misconfigurations
  Routing issues
  Powér outages
- ► Cable cuts
- ▶ Government orders



OUTAGES OUR SYSTEM OBSERVED IN 2019



LOCALIZED OUTAGE(S) ONLY

WHOLE COUNTRY SIGNIFICANTLY AFFECTED

# **IODA - Internet Outage Detection & Analysis**

Tech



https://ioda.caida.org/ioda/dashboard#view=inspect&entity=country/GM&lastView=overview&from=1641060780&until=1641429060

# **IODA - Internet Outage Detection & Analysis**



https://ioda.caida.org/ioda/dashboard#view=inspect&entity=country/GM&lastView=overview&from=1641060780&until=1641429060 Georgia

# **Kazakhstan Internet blackout** FINANCE · BITCOIN FORTUNE

# Kazakhstan internet shutdown sheds light on a big Bitcoin mining mystery

BY SHAWN TULLY

January 5, 2022 6:56 PM EST

## The hash rates for AntPool, Poolin and **Binance Pool all fell** between 12% and 16%.

The curtain just lifted on one of the many mysteries surrounding Bitcoin: How much is being produced using super-dirty coal in Kazakhstan. We knew that the Eurasian nation was a major destination for miners, and that the refugees recently expelled from China were flocking there. Still, it was difficult to establish how much of all the world's coins Kazakhstan was minting. Clouding the picture was the government's recent moves to severely restrict the mining boom that was plaguing its cities via rolling blackouts.







IODA monitoring shows Venezuela blackout starting on March 25 followed by four subsequent cascades. This electric grid is under severe stress. At 11:45pm on March 30, no sign of recovery yet. @caida\_ioda #SOSVenezuela





# Thanks

