Geopolitics of routing,

What RIPE tells us of geography and power:

a methodological approach

Loqman Salamatian and Louis Pétiniaud

UNIVERSITÉ PARIS 8 VINCENNES-SAINT-DENIS

GEODE

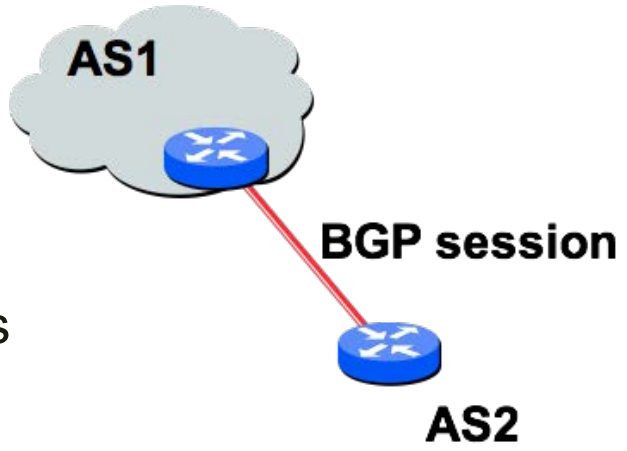
Outline

- **Representation**
  - BGP
    AS-Level Graph
    BGP Updates
    BGP AS Level Graph snapshots
  - Modeling Internet as a graph
  - Graph Embedding (I-III)
- **Topological Versus Geographical**
  - Investigating the geopolitics behind the network
  - Connectivity in the Post Soviet Space
    Investigating throught cables
  - Geopolitics of latency to Chisinau and Tiraspol
    The Reshaping of Space “from” Ukraine and Georgia
- **Borderizing in cyberspace?**
  - Is it possible to regulate Internet through routing?
  - Transferring sovereignty in Crimea through routing

Representation : BGP

BGP = **B**order **G**ateway **P**rotocol:

- is a **Policy-Based** routing protocol
- is the **de facto EGP** of today's global Internet
- is a relatively simple protocol, but configuration is complex and the entire world can see, and be impacted by, your decisions, your misconfigurations and so on...



Representation : AS-level Graph I

- Divided into Autonomous Systems
  - *Distinct regions of administrative control*
  - *Routers/links managed by a single "institution"*
  - *Service provider, company, university, …*
- Hierarchy of Autonomous Systems
  - *Large, tier-1 provider with a nationwide backbone*
  - *Medium-sized regional provider with smaller backbone*
  - *Small network run by a single company or university*
- Interaction between Autonomous Systems
  - *Internal topology is not shared between ASes*
  - *… but, neighboring ASes interact to coordinate routing*

Representation : BGP Updates

- We use bgpstream to get BGP updates from several open BGP feeds
  - All RIPE rrcxx feeds

JSON updates

–collector': 'rrc19', 'message': 'announce', 'peer': {'address': '197.157.79.173', 'asn': 37271}, 'time': 1515110408, 'fields': {'asPath': ['37271', '6939', '52320', '23106', '23106', '23106', '262700'], 'prefix': '187.102.120.0/21', 'nextHop': '197.157.79.173'},

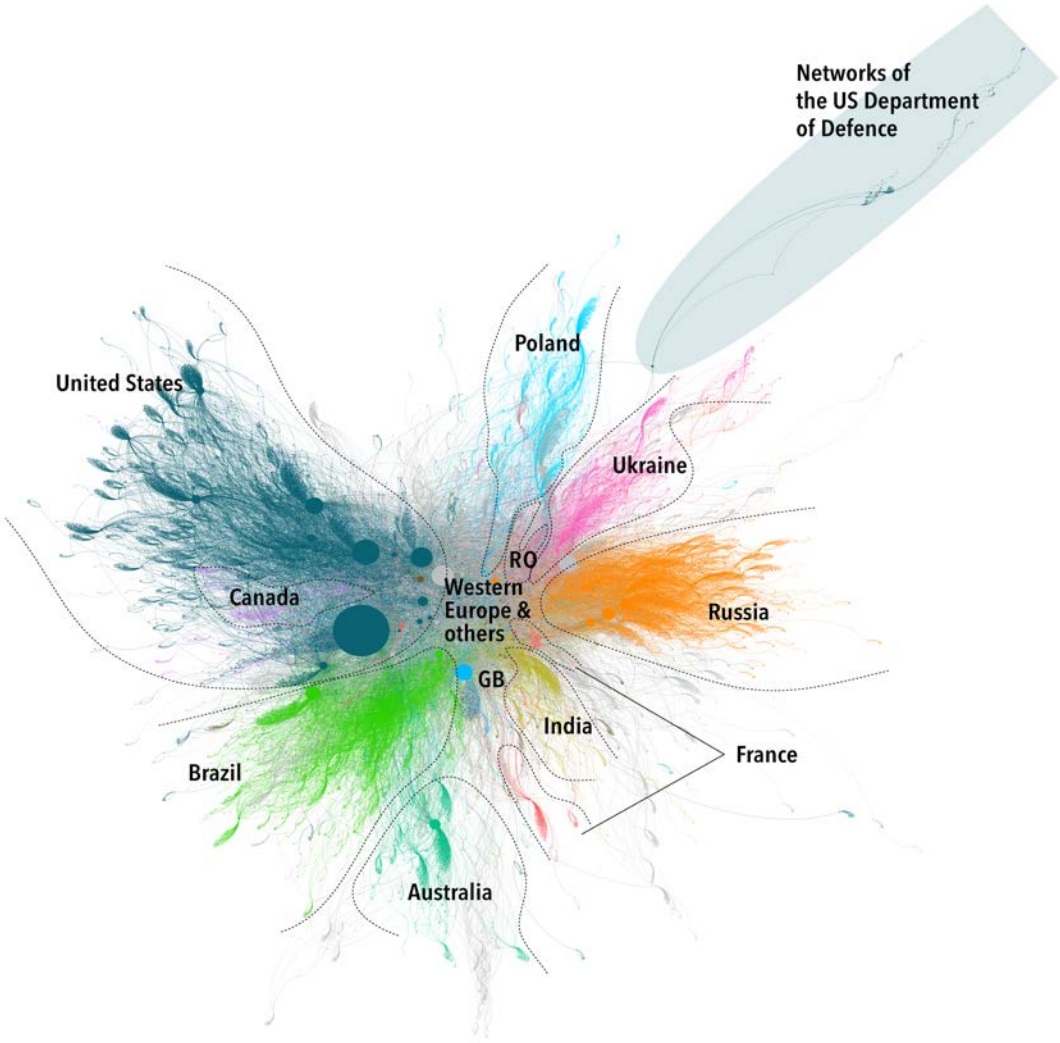
- Augmented

–'flags': {'version': 'v4', 'shortPath': ['37271', '6939', '52320', '23106', '262700'], 'geoPath': ['ZA', 'US', 'CO', 'BR', 'BR'], 'names': ['WorkonlineCommunications(Pty) Ltd', 'Hurricane Electric, Inc.', 'GlobeNetCabosSubmarinosColombia, S.A.S.', 'CemigTelecomunicaçõesSA', 'EfibraTelecom LTDA - EPP'], 'risk': 9.26246085594895e-05, 'previousPath': None, 'activePath': None, 'category': None}}

Representation : BGP AS level graph snapshots

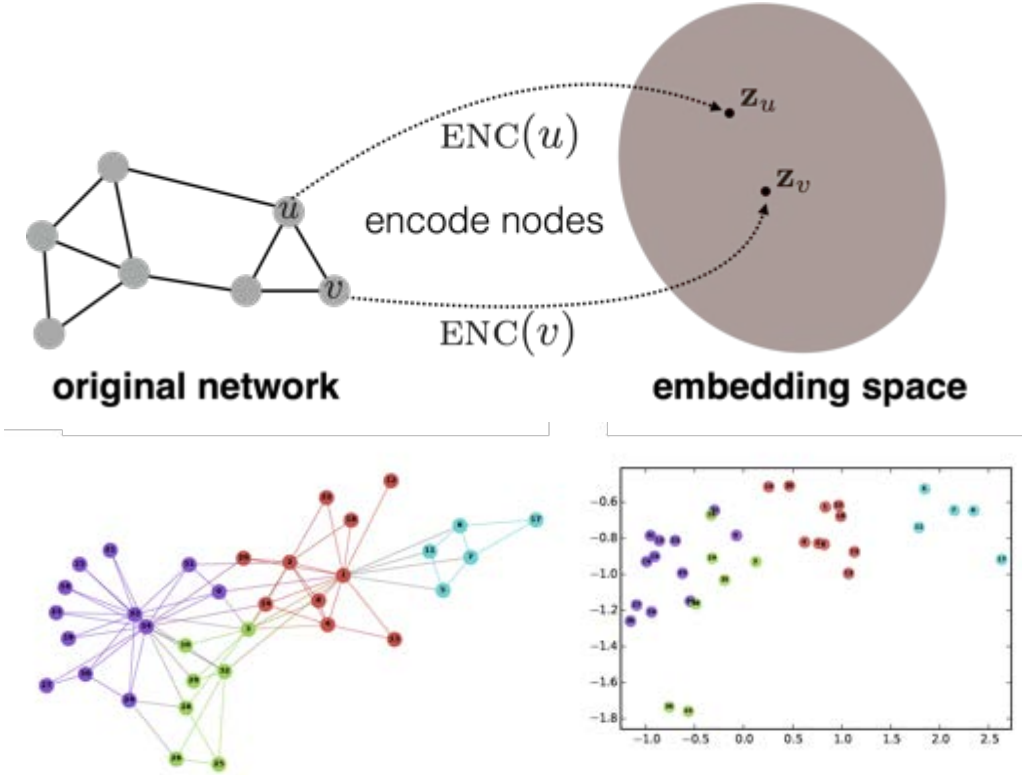
- AS paths reveal existence of link in between ASes
  - Link can be inferred to be broken when all prefixes passing by it are withdrawn.
- We infer the AS topology and updates it following BGP updates
- Caveat
  - The inferred topology is not complete as BGP announcement might be filtered
    - This is why we try to mix several feeds
- It take around one hour to get an almost full AS graph view with more than 60k AS

Representation : Modeling Internet as a graph II



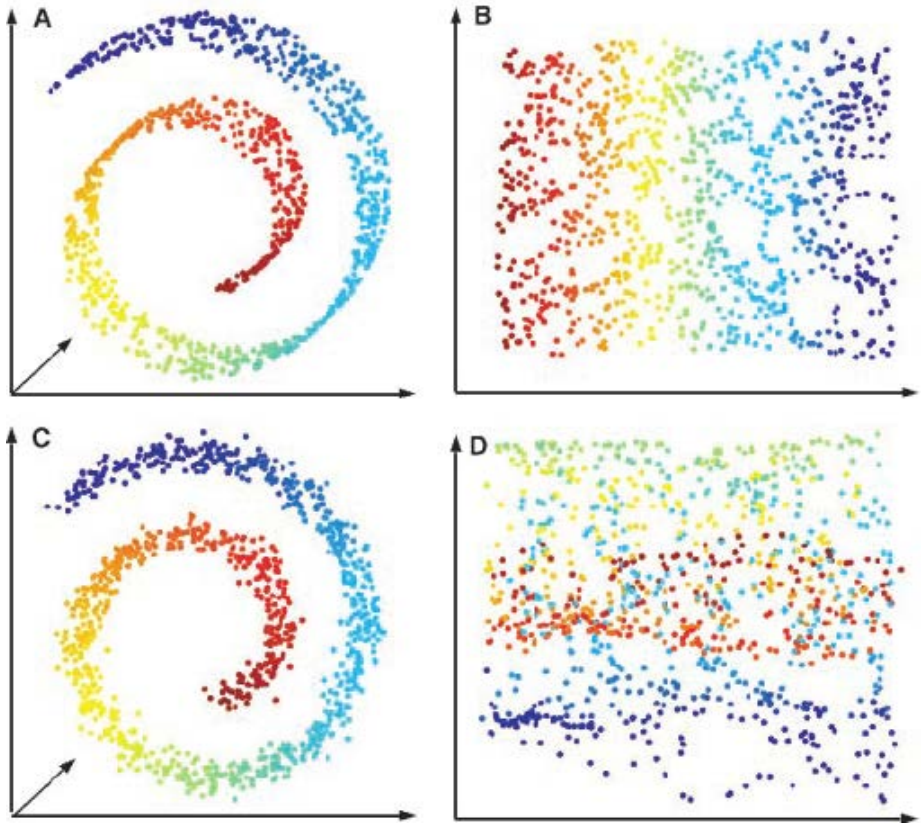
Representation : Graph Embedding I

- Goal is to encode nodes so that similarity in the embedding space approximates **similarity** in the original network



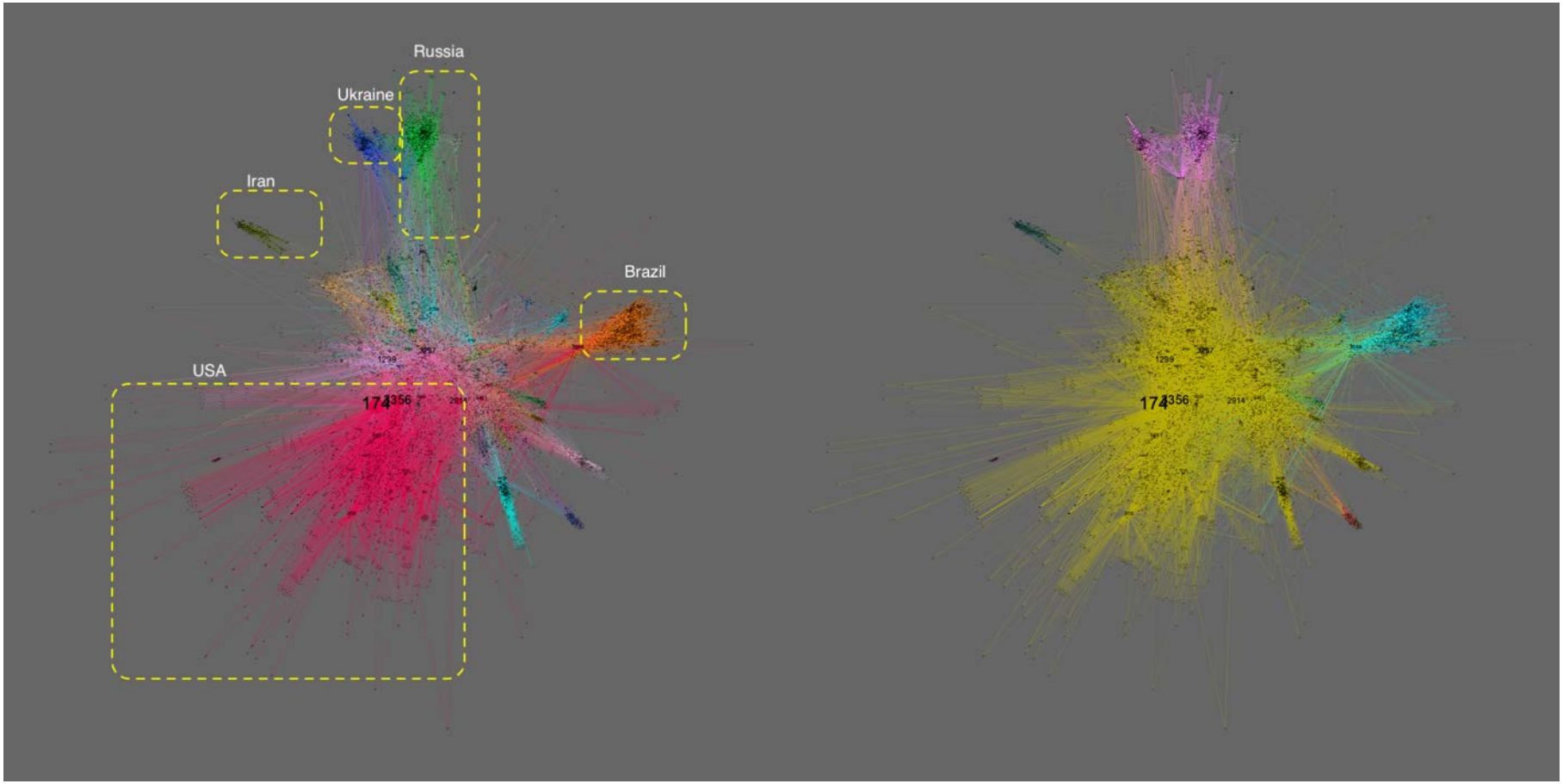
Representation : Graph Embedding II

1. Projection space
   - Generally the space basis are fixed
   - We just exchange the coordinates
2. Representation learning is the major problem
   - Can be dealt by seeing the representation as a point in a meta-space
3. Dimensionality reduction
   - Finding the good enough space for representing the information



Representation : Graph Embedding III

Spectral embedding and connectivity can be used to show zones of influence in the network



Topological versus Geographical : Investigating the geopolitics behind the network

Internet is a dual space by essence:

Geographical as a network of cable

Topological as an abstract entity of router intertwined by BGP

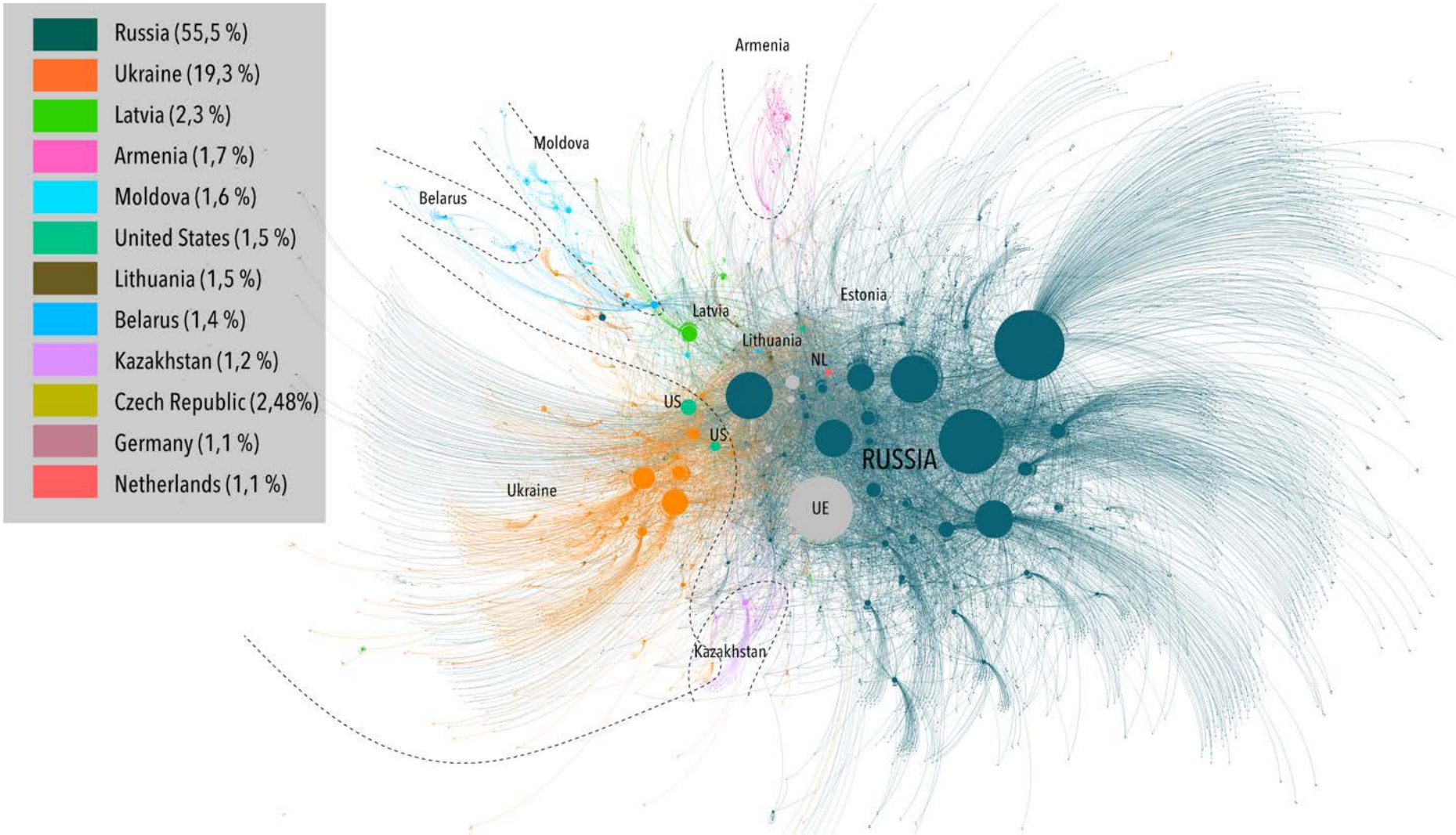
Our initial questions:

How does the geopolitical context influences the network paths?

Reciprocally, can we get a better grasp on the geopolitical context by looking at the AS-level graph?

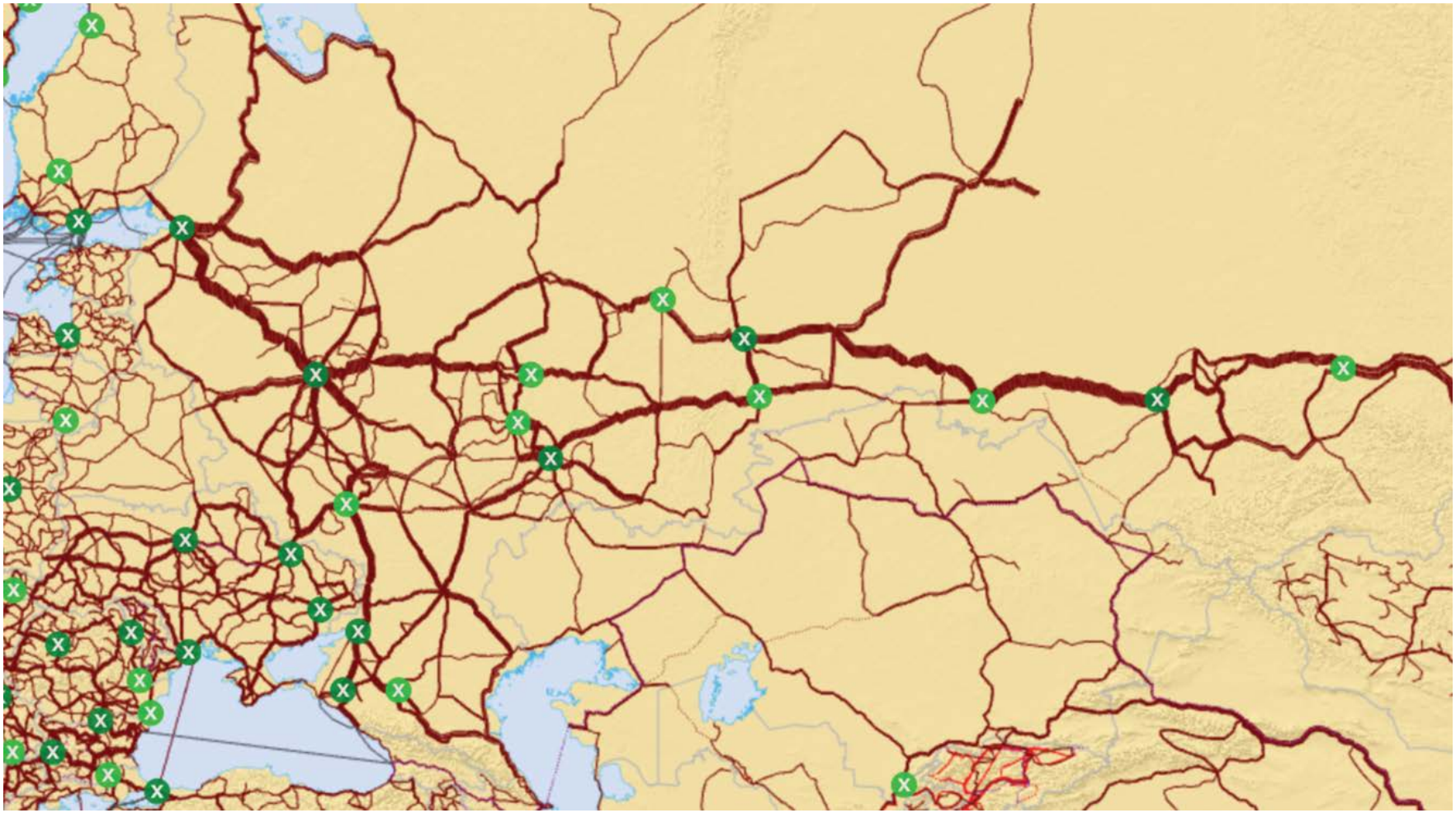
What is the exact relationship between geography and topology?

Topological versus Geographical : Connectivity in the Post Soviet space

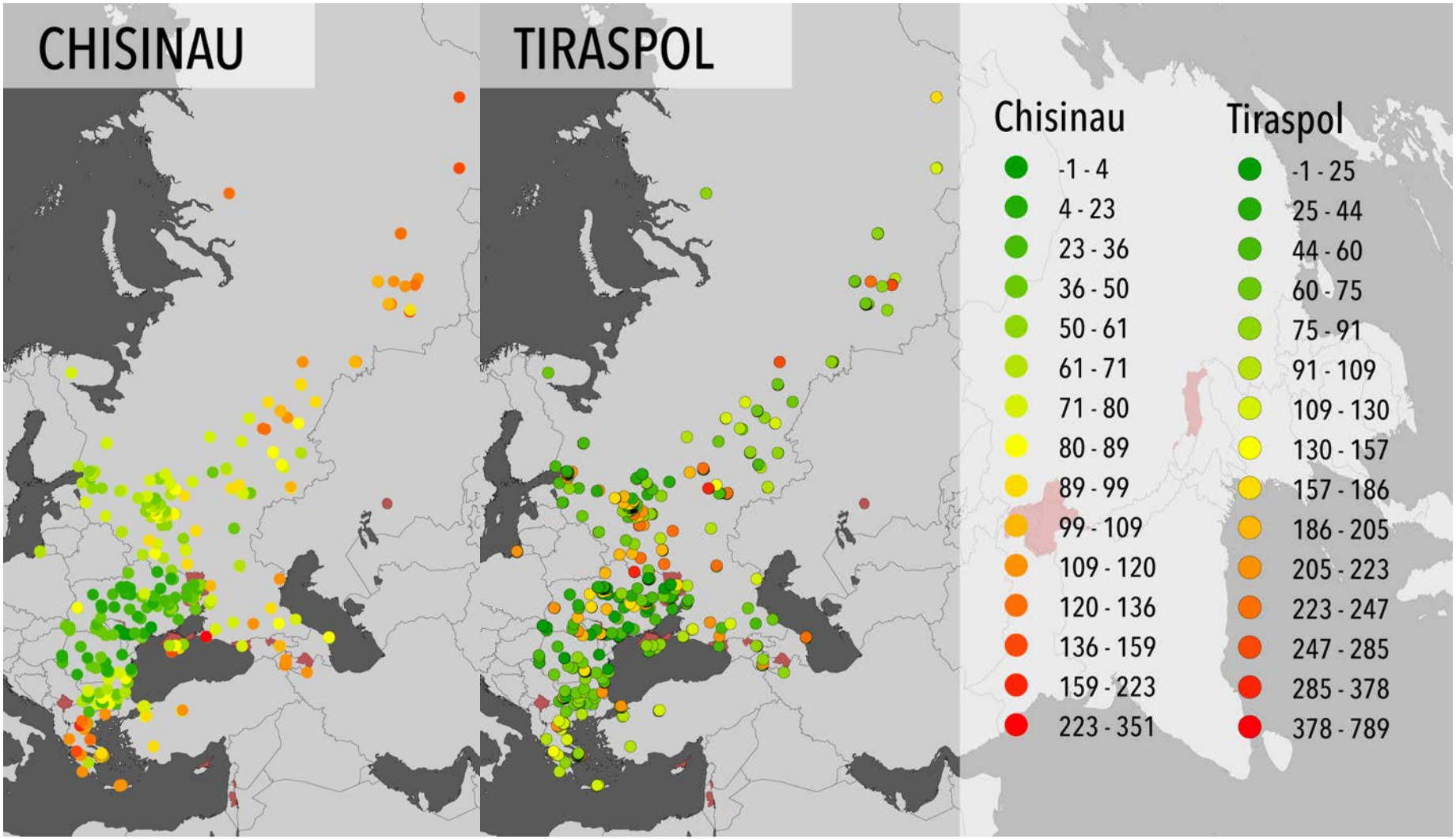


Topological versus Geographical : Investigating through cables

What is the exact relationship between geography and topology?

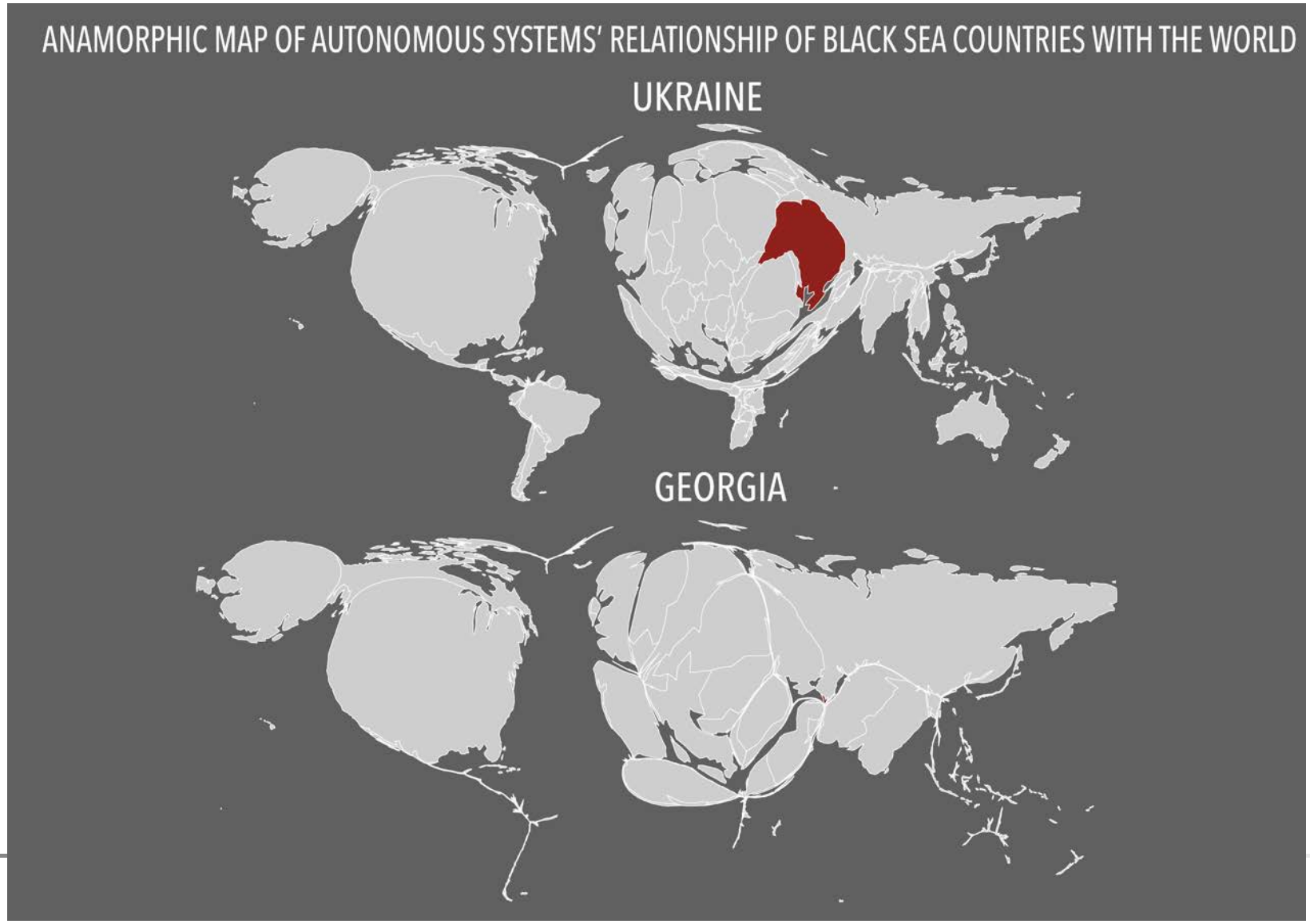


Topological versus Geographical : the Geopolitics of latency to Chisinau and Tiraspol



Topological versus Geographical: the Reshaping of Space “from” Ukraine and Georgia

How can we reshape geography according to the distance induced by the topology?



Borderizing in cyberspace (I) : Is it possible to regulate Internet through routing ?

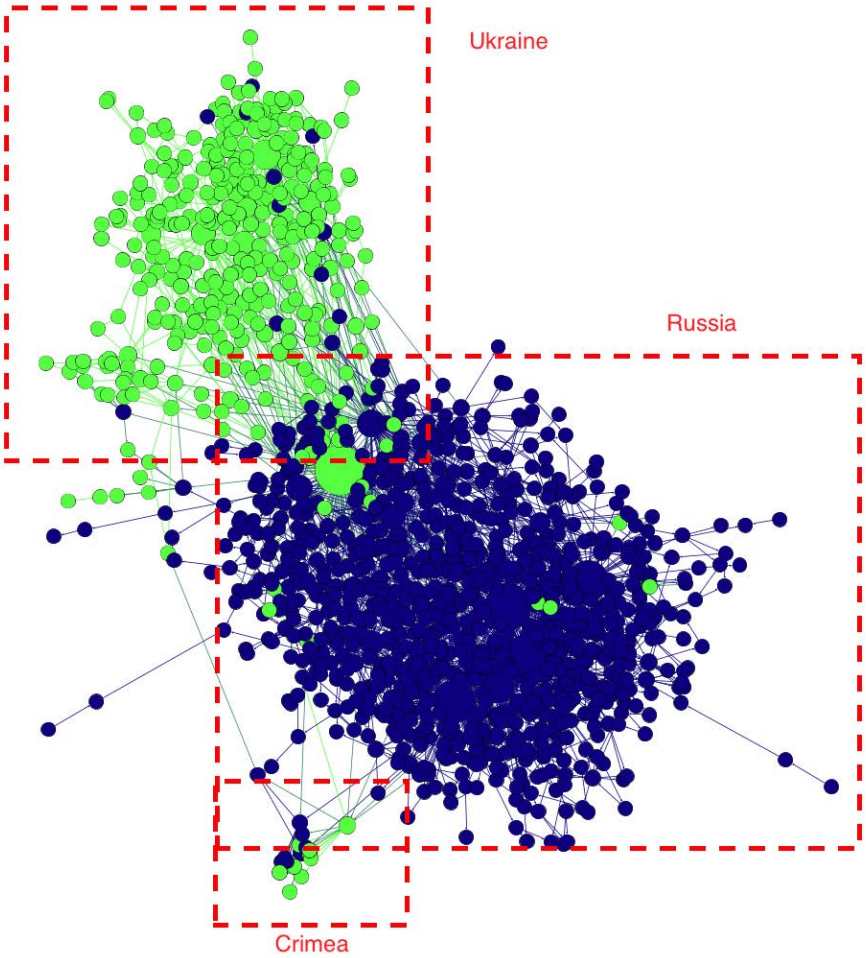
Point of control : ***the smallest set of autonomous systems that have the potential to control virtually all (90%) of the traffic within a given country***

Complexity : The network complexity metric determines the complexity of controlling who connects to the Internet within a given country

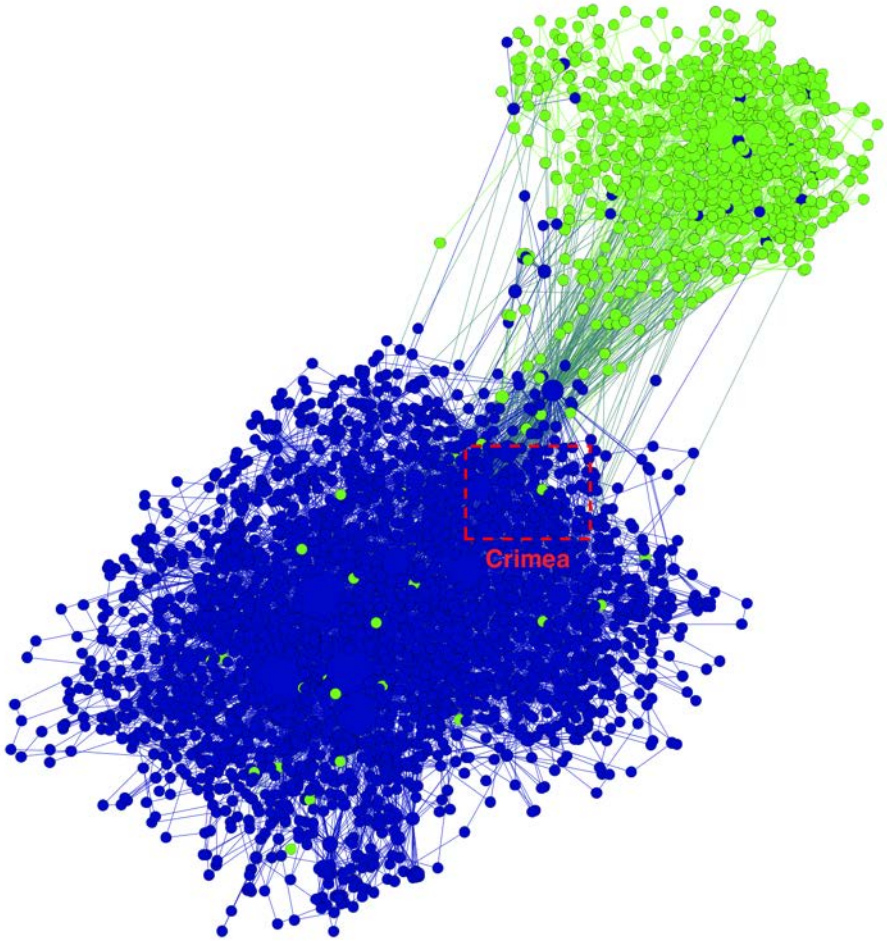
Points of Control and Network Complexity in the Post Soviet States (ranked by complexity)					
Country	Total IP addresses	Autonomous Systems	Points of Control	IP addresses/Point of Control	Complexity
Moldova	569,344	21	5	113,868	4.7
Lithuania	2,061,056	88	8	257,632	6.03
Georgia	507,648	26	3	169,216	6.35
Estonia	335,876	27	7	47,982	8.31
Kazakhstan	798,976	51	2	399,488	8.39
Azerbaijan	230,656	20	1	230,656	13.99
Kyrgyzstan	128,512	13	3	42,837	15.13
Tajikistan	34,816	6	4	8,704	18.25
Uzbekistan	138,496	20	9	15,388	19.82
Belarus	327,424	44	1	327,424	19.9
Latvia	1,314,368	156	5	262,873	24.38
Russian Federation	23,847,628	2374	36	662,434	**26**.28
Armenia	167,680	29	3	55,893	31.5
Ukraine	4,907,135	1165	47	104,407	41.57
Turkmenistan	*No data*				

Borderizing in cyberspace (II) : Transferring sovereignty in Crimea through routing

June 2017



January 2019



Conclusions :

1. **Routing and politics are deeply tied** : need for creating satisfying representations taking into account this entanglement
2. **Many unasked questions** : need for the right settings to further explore them.
3. **Multidisciplinary research is necessary** in routing analysis and cybersecurity.
4. **Need for other and/or better data sets**
    a. (Geolocated) terrestrial cables data
    b. Atlas probes located in remote regions
    c. Standardized addresses of Autonomous systems