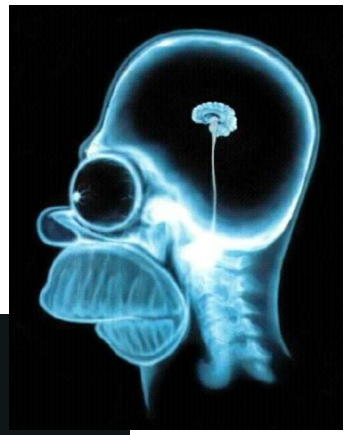
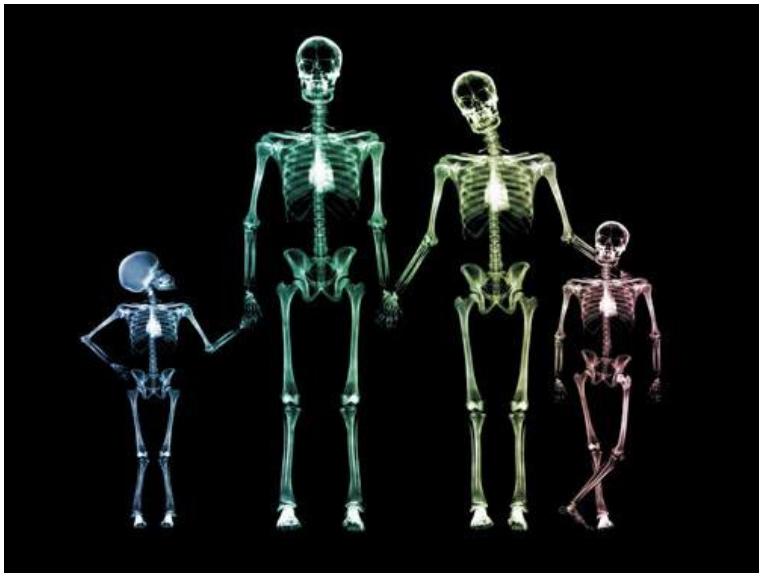
Radiography of Regional AS Interconnection





Sofía Silva Berenguer

Universidad Carlos III de Madrid

Motivation

- Understand Internet performance through info about interconnection between Autonomous Systems (ASes)
- Study regional interconnection level
- Pay attention to understudied regions. In this case, the LAC region.
- Make comparisons between regions and countries.

Goals

1) Build diagrams

- Build AS-level Internet connectivity diagrams (World and regions)
- Add local routing info from LAC
- Study impact of local info into diagrams
- Build country diagrams

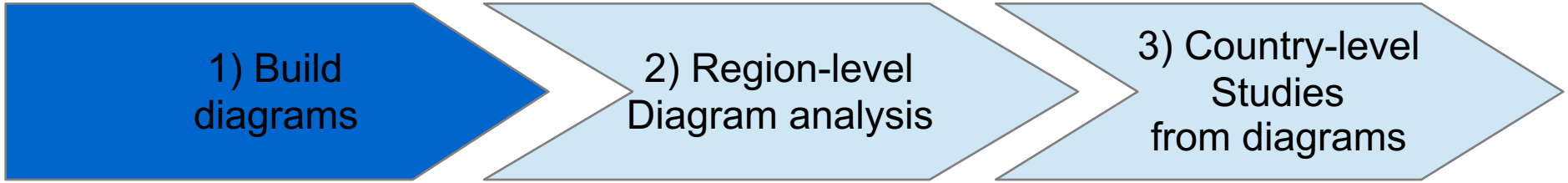
2) Region-level Diagram analysis

- Characterise diagrams
- Compare LAC diagram to other regions' diagrams

3) Country-level Studies from diagrams

- Measure interconnection level for each country
- Study IXP creation impact depending on location
- Find correlations with other indicators

Goals



- Build AS-level Internet connectivity diagrams (World and regions)
- Add local routing info from LAC
- Study impact of local info into diagrams
- Build country diagrams

- Characterise diagrams
- Compare LAC diagram to other regions' diagrams

- Measure interconnection level for each country
- Study IXP creation impact depending on location
- Find correlations with other indicators

ASes' Relationships Inference

- CAIDA's algorithm is used in order to infer relationships between ASes from routing info*
- **Base Line**: Relationships infered by CAIDA from RouteViews and RIS info from April 2015

*http://www.caida.org/data/as-relationships/

Data Sources

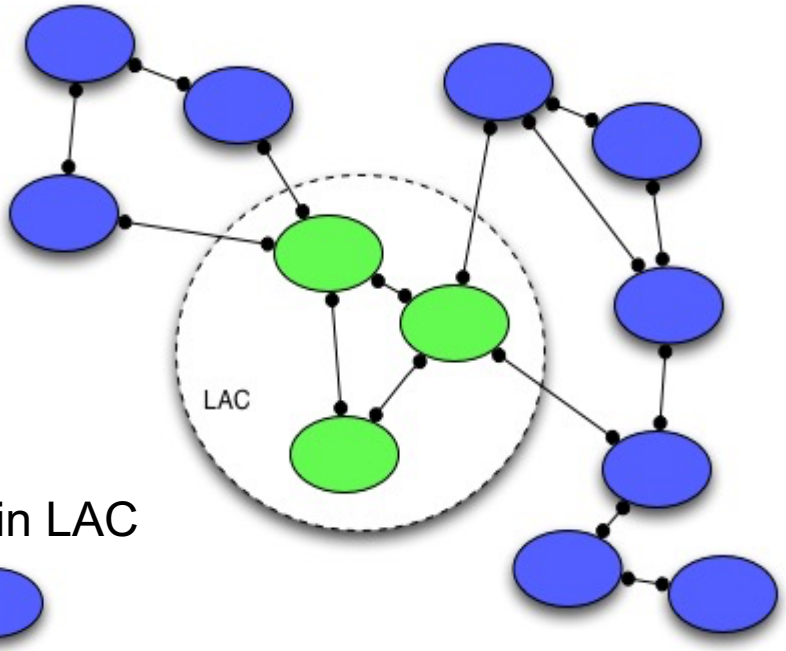
- RIS+RV Set:
  - RouteViews Project (RV) (University of Oregon)
  - Routing Information Services (RIS) (RIPE NCC)
- LAC Set:
  - Access Haiti
  - GTD Internet (Chile)
  - LACNIC
  - Packet Clearing House (pch.net) (Collectors in LAC)
  - CABASE NAPs Looking Glasses
  - PTT Metro (ix.br) Looking Glasses
  - NAP Chile Looking Glass
  - Orange Chile Looking Glass

Criteria to Define Graphs for Specific Area

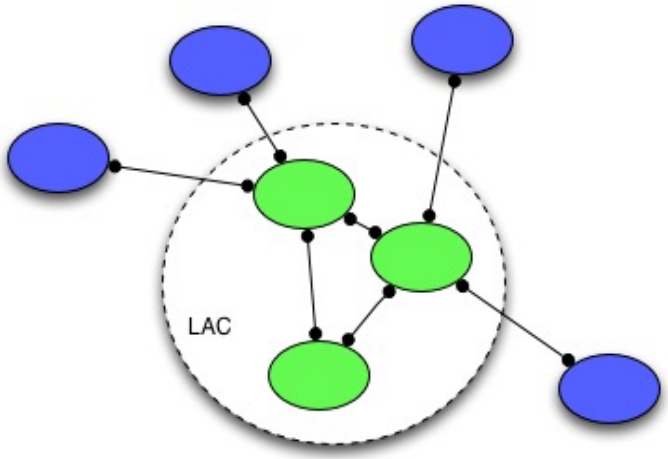
- Area → Region or country
- **Criterion 1**: Include all the relationships active in the area (at least one of the ASes is active in the area) and all the ASes involved in these relationships.
- **Criterion 2**: Include all the ASes active in the area and all the relationships in which they are involved.
  - *To keep this presentation short, we will only show results for Criterion 2.*

- RIPEstat API was used in order to geolocate ASes. An AS is active in a country if it is announcing at least one prefix that is geolocated to that country.

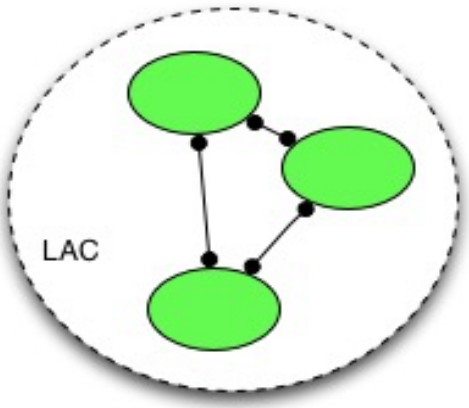
Criteria to Define Graphs for Specific Area



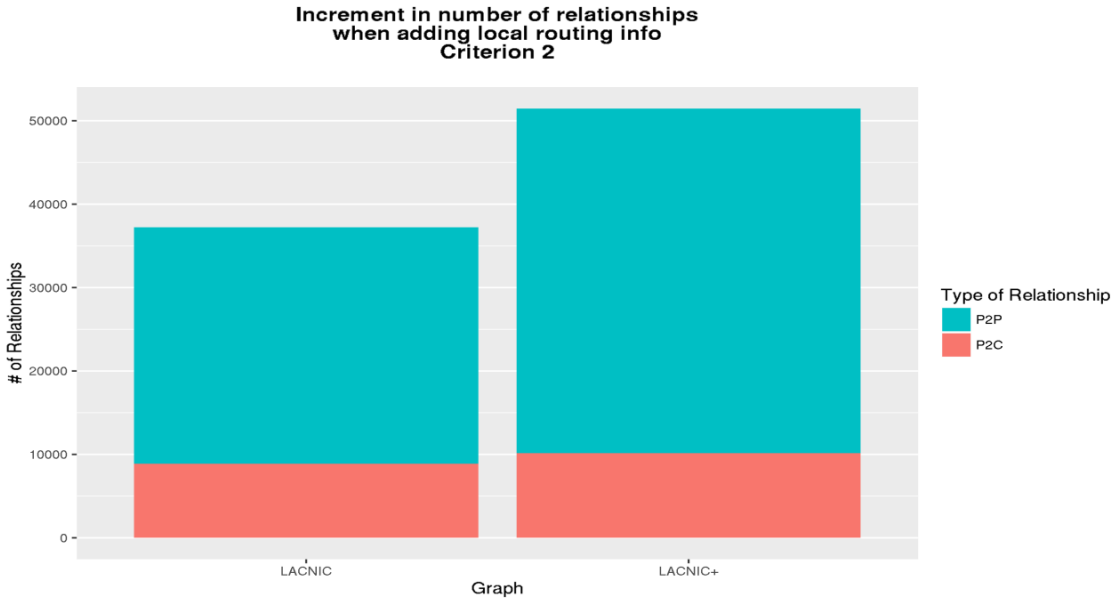
Criterion 1
Relationships active in LAC



Criterion 2
ASes active in LAC



Adding Local Routing Info to LAC Graph



- 37,234 relationships inferred from RIS+RV set
- 51,479 relationships inferred from RIS+RV+LAC set (38.3 % increment)
  - 12,954 additional P2P relationships (90.9 %)
  - 1,291 additional P2C relationships (9.1 %)

Goals

- Build AS-level Internet connectivity diagrams (World and regions)
- Add local routing info from LAC
- Study impact of local info into diagrams
- Build country diagrams

- Characterise diagrams
- Compare LAC diagram to other regions' diagrams

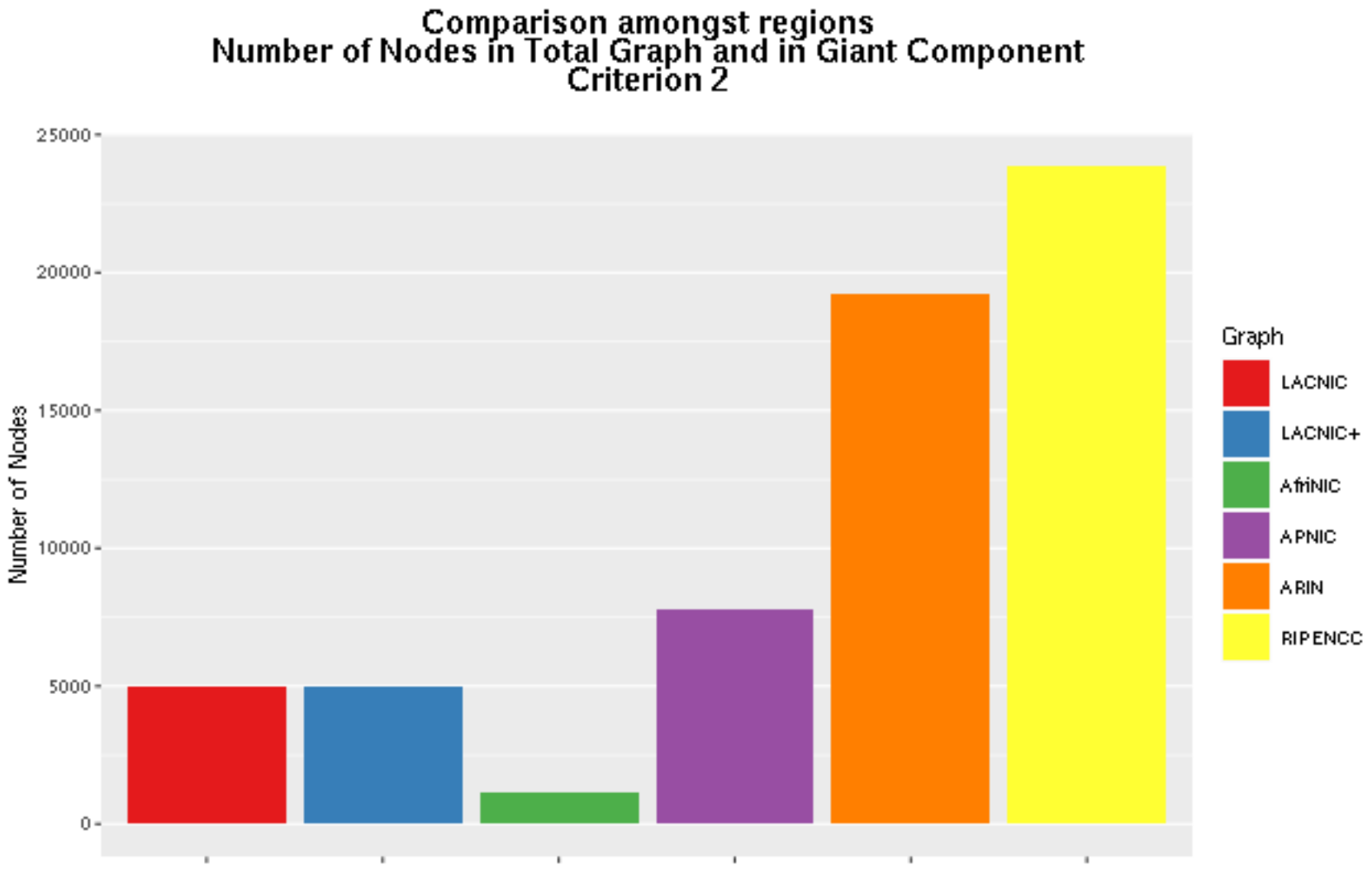
- Measure interconnection level for each country
- Study IXP creation impact depending on location
- Find correlations with other indicators

Graph Characterisation*

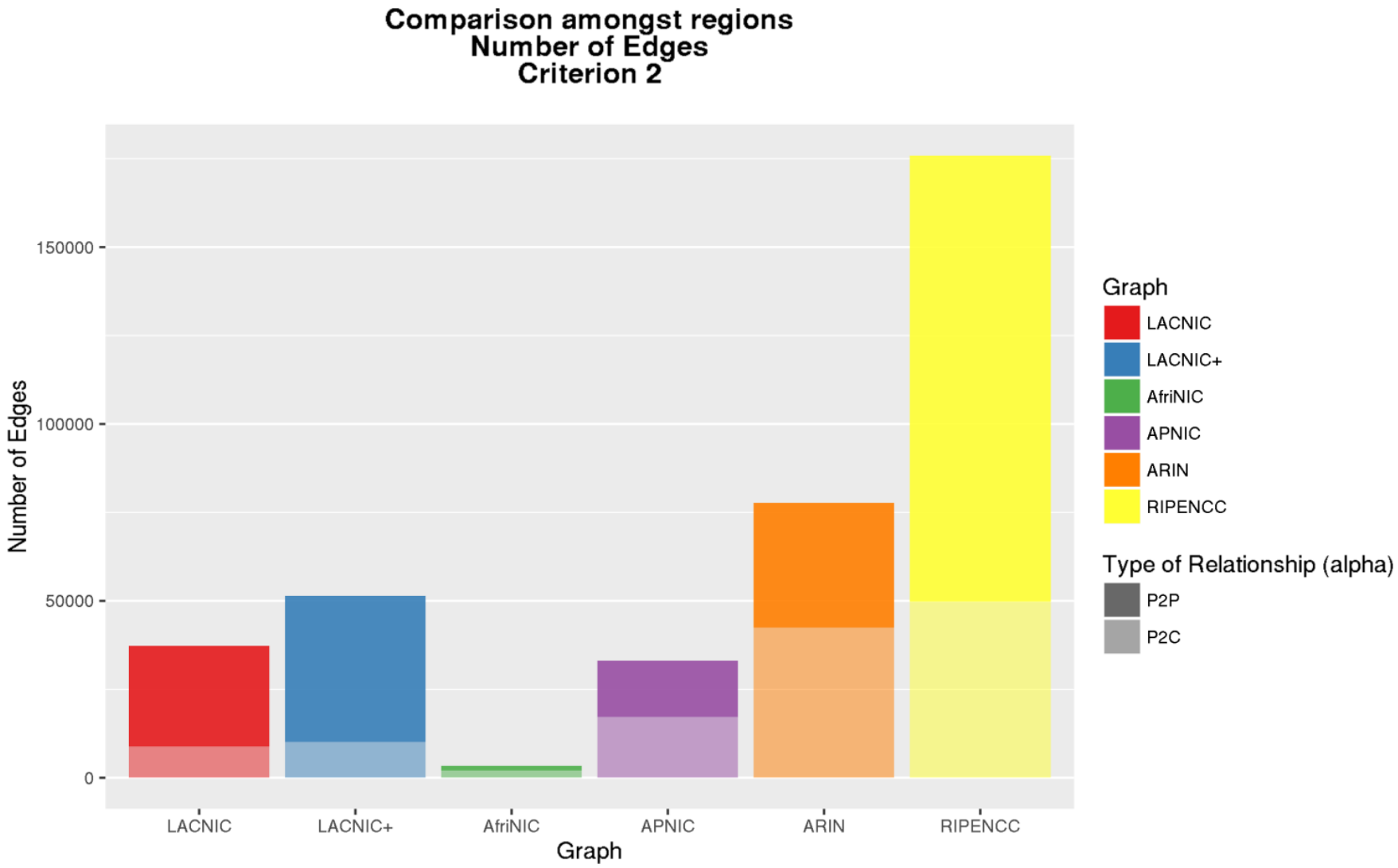
- **Order** *(# of Nodes (ASes))*
- **Size** *(# of Edges (Relationships))*
- **Giant Component and Disconnected Nodes**
- **Degree Distribution** → *Degree of a node is the number of relationships the nodes is involved in*
- **Distance Distribution** → *Distance as the minimum number of AS-hops between two ASes*
- **Avg Clustering Coefficient vs Degree** → *Avg of the clustering coefficients of the nodes with degree k*

* Based on book Network Science (Barabasi) (http://barabasi.com/networksciencebook/)

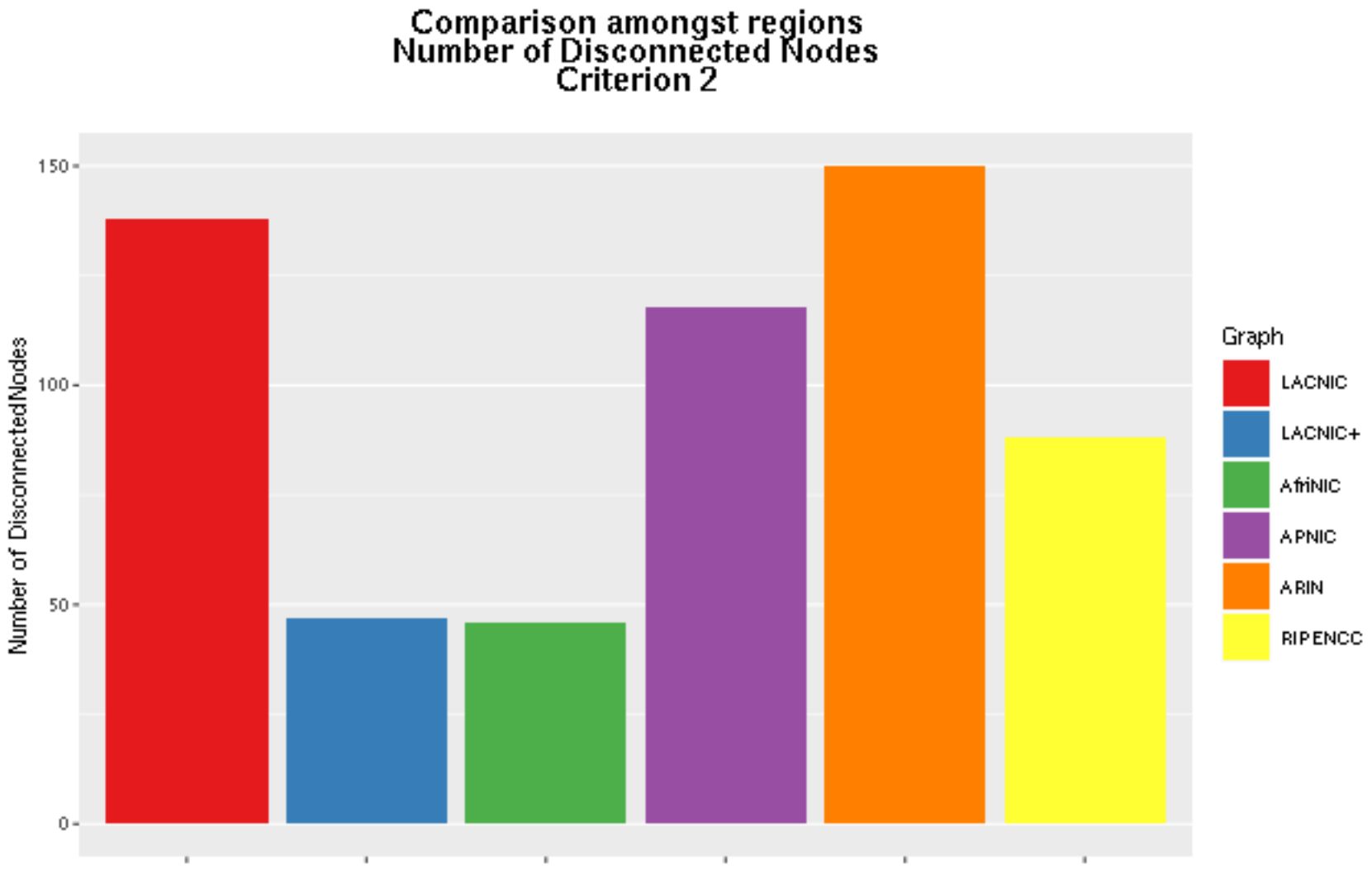
of Nodes


Comparison amongst regions
Number of Nodes in Total Graph and in Giant Component
Criterion 2
Number of Nodes
25000
20000
15000
10000
5000
0
Graph
LACNIC
LACNIC+
AfriNIC
APNIC
ARIN
RIPENCC

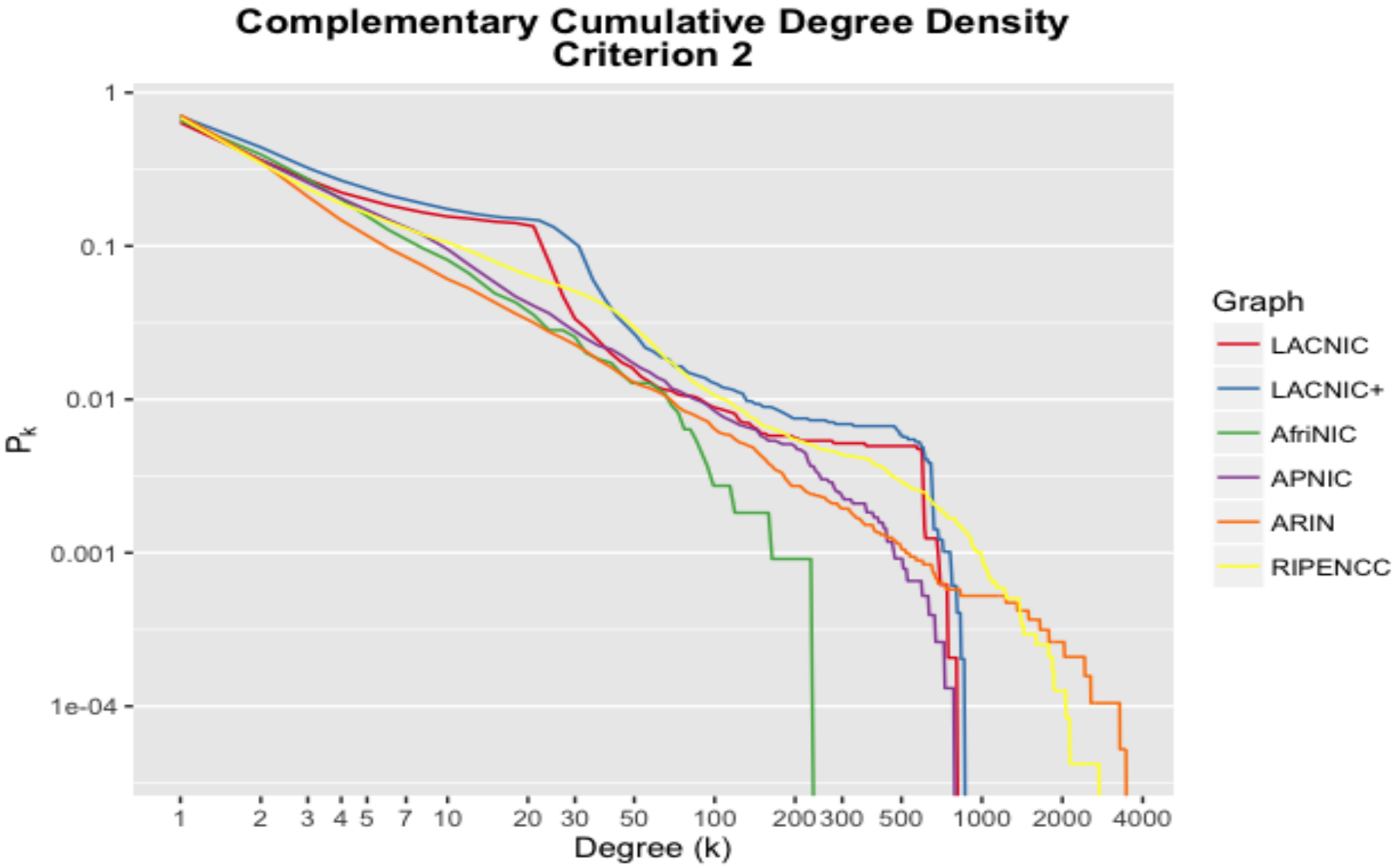
of Edges


Comparison amongst regions
Number of Edges
Criterion 2
Number of Edges
150000
100000
50000
0
LACNIC
LACNIC+
AfriNIC
APNIC
ARIN
RIPENCC
Graph
Graph
LACNIC
LACNIC+
AfriNIC
APNIC
ARIN
RIPENCC
Type of Relationship (alpha)
P2P
P2C

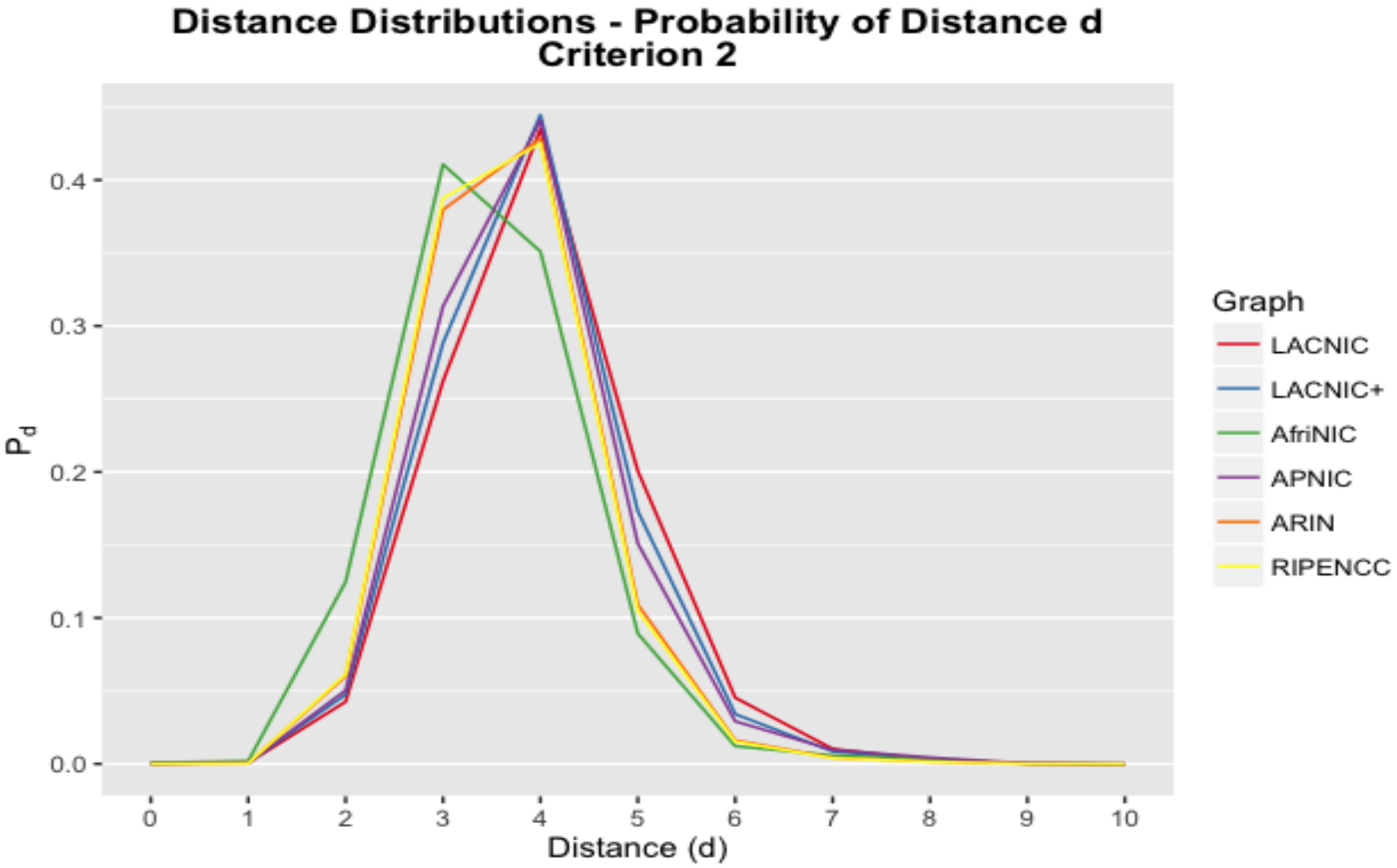
Disconnected Nodes



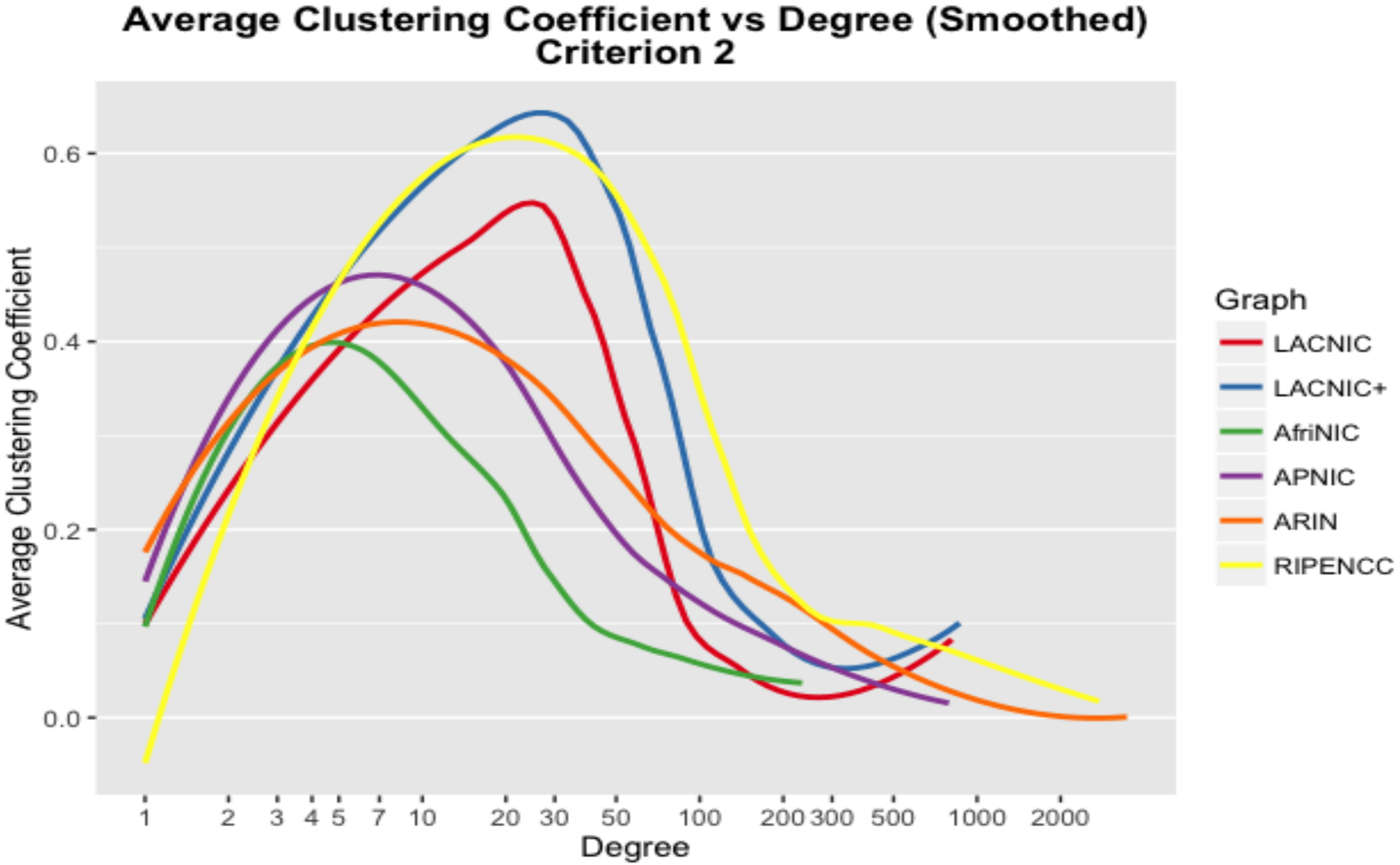
Degree Distribution



Distance Distribution



Avg CC vs Degree



Observations

- LACNIC is the 2nd smallest in terms of # of nodes but is bigger than APNIC in terms of # of edges.
- **Disconnected Nodes:**
  - Probably misgeolocated nodes (legacy, reserved, unassigned)
  - Huge decrement for LACNIC when adding local routing info, because new relationships between ASes active in the LAC region are discovered.
- **Degree Distribution**
  - Approx. Power-laws
  - Interesting peaks for LACNIC (around k=20 and k=500). All the ASes with k~=500 are active in Brazil and more than 90 % of them were assigned to Brazil.

Observations (Cont.)

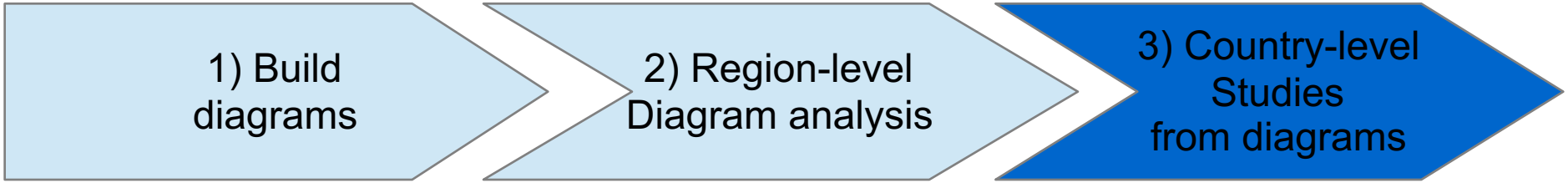
.Distance Distribution

- Except for AfriNIC, the peak occurs at d = 4 (more than 40 % of the paths) for all the regions and the second most probable distance is d=3. It is the other way around for AfriNIC (more than 40 % of the paths are 3 AS-hops long).

.Avg CC vs Degree

- Again the LACNIC graphs show increments around k=20 and k=500 (Brazil-effect)
- All the graphs are decrescent after a peak (ASes with high k are usually in sparse neighbourhoods while there's a medium layer (k corresponding to peak) at which the local neighbourhoods are dense.

Goals



- Build AS-level Internet connectivity diagrams (World and regions)
- Add local routing info from LAC
- Study impact of local info into diagrams
- Build country diagrams

- Characterise diagrams
- Compare LAC diagram to other regions' diagrams

- Measure interconnection level for each country
- Study IXP creation impact depending on location
- Find correlations with other indicators

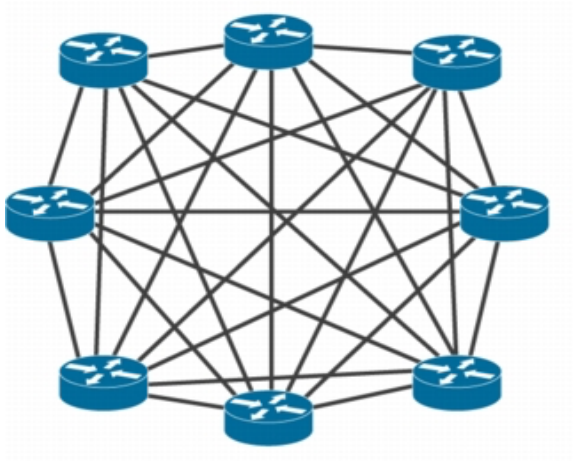
Avg Degree

as a means of measuring interconnection level?

Median Degree

as a means of measuring interconnection level?

Simulation of Interconnection at IXPs



- **Dominican Republic, Guatemala and Mexico**

- **Assumptions:**
  - Mandatory Multilateral Peering
  - ASes that are already connected to other IXPs in the region and that are active in the country, will get connected to the simulated IXP.
  - Google and Akamai get connected to the IXP.
  - The 1st 10% biggest ASes (with highest Degree) active in the country get connected to the IXP.

Impact of the “new” IXPs

	IXP DO	IXP GT	IXP MX
Region-level			
New relationships	8	36	594
Δ # of relationships	0.03 %	0.12 %	1.93 %
Δ Avg Degree	0.01%	0.1 %	1.91 %
Δ Med Degree	0 %	0 %	0 %
Country-level			
New relationships	12	39	594
Δ # of relationships	17.65 %	57.35 %	94.89 %
Δ Avg Degree	12.04 %	51.42 %	94.19 %
Δ Med Degree	0 %	0 %	0 %

Future Work

- Find correlations of national graph metrics with other indicators (economic (E.g. GDP) and transport (airports, flights, etc.))
- Check correlation between interconnection level and delay between countries (LACNIC's SIMON project)
- Analize outages impact

Conclusions

- There's room for improvement in terms of interconnection in the LAC region.
- Having local routing info is highly important.
- With better diagrams we can better understand the regional Internet performance and find critical aspects to work on.
- Lots of interesting studies could be done.
- These studies could help finding incentives for governments and other entities to promote and facilitate the creation of IXPs and for the big ASes to get connected to those IXPs.

Questions?
Comments?
Feedback?
Suggestions?
Ideas?
Gossip, jokes, etc.? :)

ssilva@it.uc3m.es

Open Questions

- **Criteria to Define Graphs for Specific Area**
  - *Which criterion do you think should be used?*

- **Graph Characterization**
  - *Should we also consider Betweenness Centrality or any other centrality metrics?*
  - *Should these node metrics be computed for the World graph and considered a charateristic of the node no matter which area the graph is restricted to? Or should they be computed for each regional or country-wide graph?*

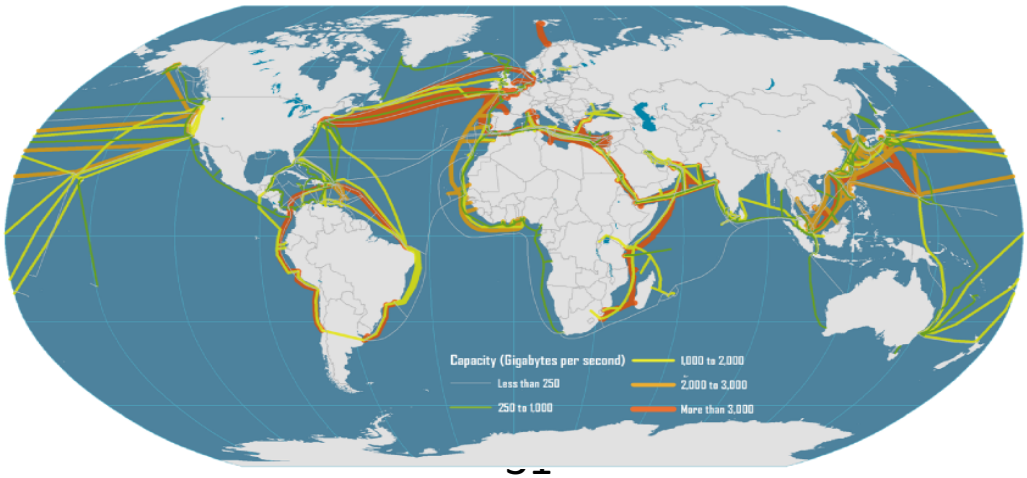
- **In general:**
  - *Is there anything else we can conclude from these metrics?*


thank you
danke
謝謝
ngiyabonga
рахмат
teşekkür ederim
спасибо
gracias
dank je
tapadh leat
obrigado
bedankt
dziękuję
merci
go raibh maith agat
sukriya
kop khun krap
terima kasih
grazie
arigatô
dakujem
takk
tack
barka
welalin
mersi
spas
Баярлалаа
faafetai lava
vinaka
blagodaram
misaotra
matondo
paldies
grazzi
mahalo
kiitos
dankie
dhanyavad
nanni
nandri
mauruuru
hvala
köszönöm
bayarlalaa
grăçie
enkosi
akun
dankon
ačiū
djiere dieuf
tau
mochchakkeram
дякую
mamnun
asante
manana
obrigada
tenki
murakoze
chokrane
хвала
chnorakaloutioun
gratias ago
gràcies
sulpáy
taiku
sobodi
děkuji
sagolun
mèsi
didi madloba
kam sah hamnida
najis tuke
rahmat
ありがとう
tanemirt
rahmet
তোমাকে ধন্যবাদ
감사합니다
xièxie
ευχαριστώ
diolch
dhanyavadagalu
shukriya
mercè
мерси
trugarez

Backup Slides

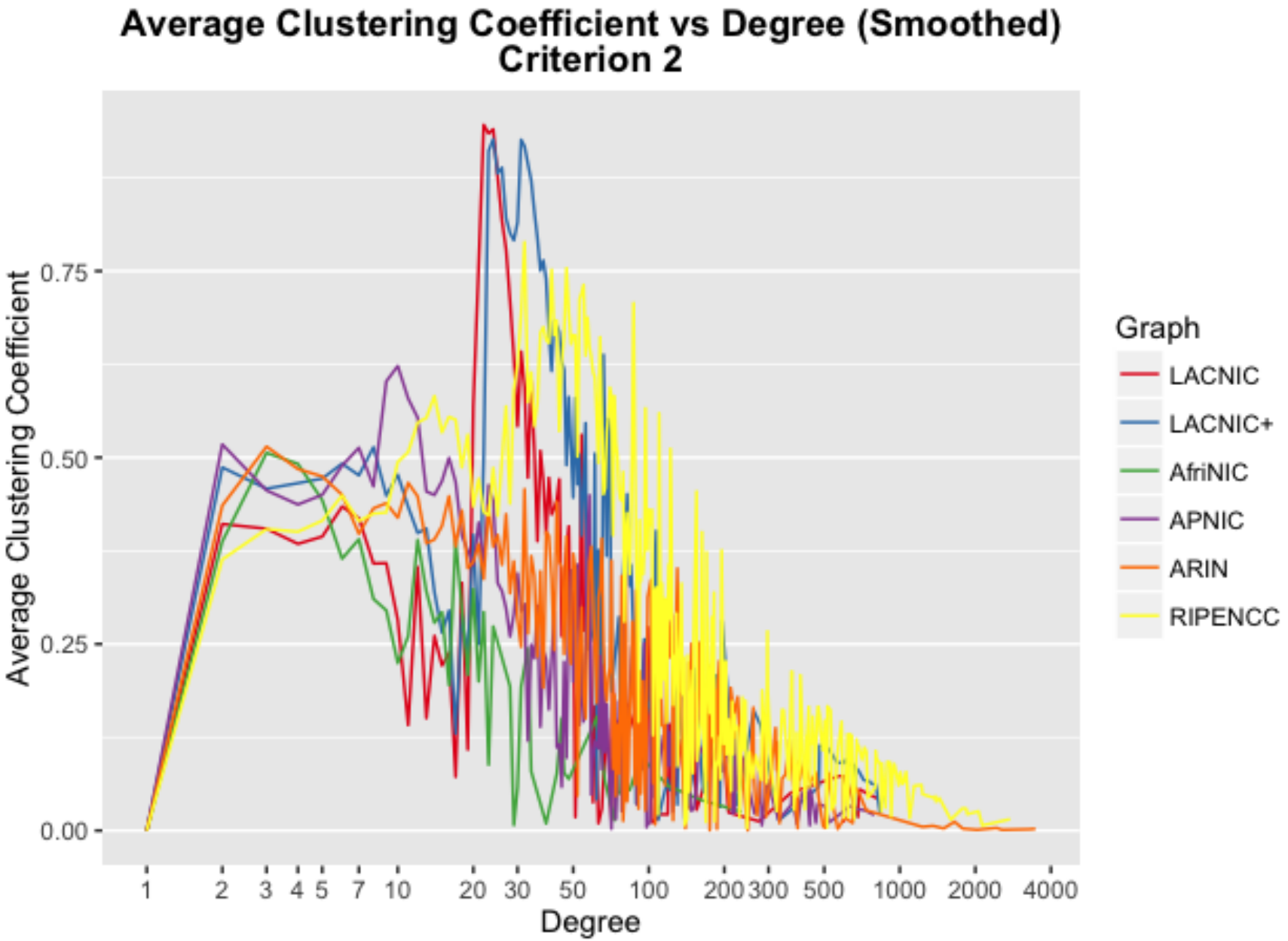
Why LAC?

- Poorly interconnected region. Few local traffic exchange.
- Geography issues → few cables.
- Few local infrastructure.
- Strong dependence in northamerican infrastructure.
- Few routing info collectors



Importance of Improving Interconnection in the Region

- Lower co$ts!
- Better Internet performance in the region (Lower delay)
- More security and robustness
- More possibilities for innovation development (New local businesses could appear)


Average Clustering Coefficient vs Degree (Smoothed)
Criterion 2
Average Clustering Coefficient
0.75
0.50
0.25
0.00
1
2
3
4
5
7
10
20
30
50
100
200
300
500
1000
2000
4000
Degree
Graph
LACNIC
LACNIC+
AfriNIC
APNIC
ARIN
RIPENCC

Thanks to...

- Andra Lutu (Background, source code and much more things)
- CAIDA (ASes' relationships inference algorithm)
- Juan Camilo Cardona (IMDEA Networks) (Algorithm to process “show ip bgp” outputs)
- Routing Data Sources