



CSCI 4460/5460 **NETWORK OPERATION & DEFENSE**

Fall 2022

LAB #3: Network Measurement and Analysis

DUE: October 31, 11:59pm

GOALS

In this assignment, we will learn details about Degree Centrality in graph theory for network analysis.

RESOURCES & ASSUMPTIONS

The plan:

We have a dataset contains AS relationships, downloaded from [1]. The dataset date is 2022-10-01. For more information about how the dataset is generated, you can check [2].

You can download the dataset from the following link:

<https://ayasinnur.com/uno4460/ASRelationship-2022-10-01.txt>

In case you are not able to download txt file (some browsers open txt file instead of downloading), you can download the following compressed file.

<https://ayasinnur.com/uno4460/ASRelationship-2022-10-01.rar>

The dataset contains p2p and p2c relationships. It is separated by "|".

The format is: ASNumber1 | ASNumber2 | Relationship | Source

<provider-as>|<customer-as>|-1|<source>

<peer-as>|<peer-as>|0|<source>

E.g.

1|5467|0|bgp

ASN = 1 has a p2p relationship with ASN = 5467.

1|5|-1|bgp

ASN1 is provider of ASN5, AS5 is customer of ASN1.

OBJECTIVES:

1-) Show how many unique ASN in the dataset.

2-) Show how many unique links in the dataset.

Remember that AS1-AS2 link also means AS2-AS1 link. However, you need to put either AS1-AS2 in your list or AS2-AS1, not both.

Hint: You can use HashSet and AS1-AS2 and AS2-AS1 in the set. Your final answer will be the $(\text{setSize}/2)$

e.g. E.g. Assume that the dataset has following lines:

1|5467|0|bgp

5|1|0|bgp

HashSet.add(1-5467)

HashSet.add(5467-1)

HashSet.add(5-1)

HashSet.add(1-5)

UniqueConnectionSize = hashSet.size / 2

3-) Show how many unique p2p and p2c relations in the dataset

e.g. 1000 p2p relations and 2000 p2c relations.

4-) Find top-10 largest ASN in terms of connection size (Node degree distribution).

You can ignore the business relations since it asks all connections. Remember that AS1-AS2 link also means AS2-AS1 link.

E.g. Assume that the dataset has following lines:

1|5467|0|bgp

5|1|0|bgp

2|5|0|bgp

1|8|0|bgp

ASN1 degree is 3 (1-5467, 5-1, and 1-8)

AS5467 degree is 1 (1-5467)

ASN2 degree is 1 (2-5)

ASN5 degree is 2 (5-1, 2-5)

ASN8 degree is 1 (1-8)

5-) Find top-10 largest Provider ASN in terms of connection size (Node degree distribution).

E.g. <provider-as>|<customer-as>|-1|<source>

1|5467|-1|bgp

5|1|0|bgp

2|5|-1|bgp

1|8|0|bgp

ASN1 provider degree is 1 (1-5467)

ASN2 provider degree is 1 (2-5)

5-1 and 1-8 are peer to peer, ignore them.

6-) Find top-10 largest peer to peer ASN in terms of connection size (Node degree distribution).

E.g. <peer-as>|<peer-as>|0|<source>

1|5467|-1|bgp

5|1|0|bgp

2|5|-1|bgp

1|8|0|bgp

ASN1 p2p degree is 2 (5-1, 1-8)

ASN5 provider degree is 1 (5-1)

ASN8 provider degree is 1 (1-8)

1-5467 and 2-5 are provider to customer relationship, ignore them.

7-) Show all tier-1 ASNs:

Tier-1 ASN: They do not have any provider, they have peers and customers.

E.g.

E.g. <provider-as>|<customer-as>|-1|<source>

1|5467|-1|bgp

5|1|0|bgp

Any ASN in the second spot of a business relation with -1 cannot be a tier-1 AS. For example, 5467 is not tier-1 AS. ASN1 can be, if it doesn't have any provider. Note that, p2p relations are ok.

Check all dataset and find only ASNs that are not customers.

8-) Show how many stub AS in the dataset.

Stub-AS: they may have providers and peers but they do not have customers.

What to Deliver:

Write a scientific report. Your report should have an Introduction section (2-3 paragraphs), related work, experimental results section, and conclusion section. Also, compress your project folder containing your source code and submit that as well.

Guidelines:**Introduction:**

1 paragraph for overall explanation. E.g. What is the Internet, what is an autonomous system, what is an ISP etc.

1 paragraph for AS business relations, tier classification of ASes (tier-1, tier-2, stub AS etc.).

1 paragraph for shortly explaining your motivation, objectives, and general explanation about your report.

Related work:

Discuss other possible graph metrics that we can use. E.g. besides degree centrality, closeness centrality, Betweenness centrality, Eigenvector centrality, PageRank centrality etc. are highly used techniques in network analysis.

Experimental section:

You need to discuss all details about your findings. Did you find anything interesting? E.g. You will see some ISPs have larger degree but they are not listed as tier-1 AS.

For each step (1 to 8), print some figures, tables etc. and discuss about them. Do NOT just give a figure, you also need to explain and discuss about it.

Conclusion:

1-2 paragraph to conclude your report.

Note that, this is just a minimum guideline. Your paper may have more sections, more details, etc. Also, besides 8 experiments, you can show implement and discuss more experiments.

BONUS OPTION:

Top 3 paper will receive 20 bonus points from this lab work.

Rules:

- You may search freely online reference materials – man pages, stack overflow, etc. However, you may not ask anyone for specific help (related to this lab) online, or offline, other than the instructor.
- If you use an information from any source, you need to give a reference. Do NOT copy paste any information from anywhere (copy-paste is called plagiarism even if you give a reference). However, you can paraphrase it and give a reference to that source.
- This is an individual assignment; all work must be yours.
- Your assignment will not be graded without a report.
- The submission deadline is a hard one. No late submissions.
- You can implement your code in any programming language. (Preferably Java)

References:

[1] <https://publicdata.caida.org/datasets/as-relationships/serial-2/>

[2] Luckie, Matthew, et al. "AS relationships, customer cones, and validation." Proceedings of the 2013 conference on Internet measurement conference. 2013.

https://www.caida.org/catalog/papers/2013_asrank/asrank.pdf

You can also read the following papers to get more information about autonomous systems and their analysis.

[3] M. E. Tozal, "Autonomous system ranking by topological characteristics: A comparative study." 2017 Annual IEEE International Systems Conference (SysCon). IEEE, 2017.

<https://people.cmix.louisiana.edu/tozal/publications/systems2017-2-tozal.pdf>

[4] Abdullah Yasin Nur and M. E. Tozal, "Identifying Critical Autonomous Systems in the Internet", Springer Journal of Supercomputing, SI: Cyber Threats against Critical Infrastructure and their Countermeasures, Volume 74, Issue 10, pp. 4965-4985, 2018

<https://ayasinnur.com/wp-content/uploads/supe2018.pdf>

[5] Abdullah Yasin Nur, "Analysis of Autonomous System Level Internet Topology Graphs and Multigraphs", IEEE International Symposium on Networks, Computers and Communications (ISNCC), 2021

<https://ayasinnur.com/wp-content/uploads/ASLandASMLComp.pdf>