Project 2

ECE 578

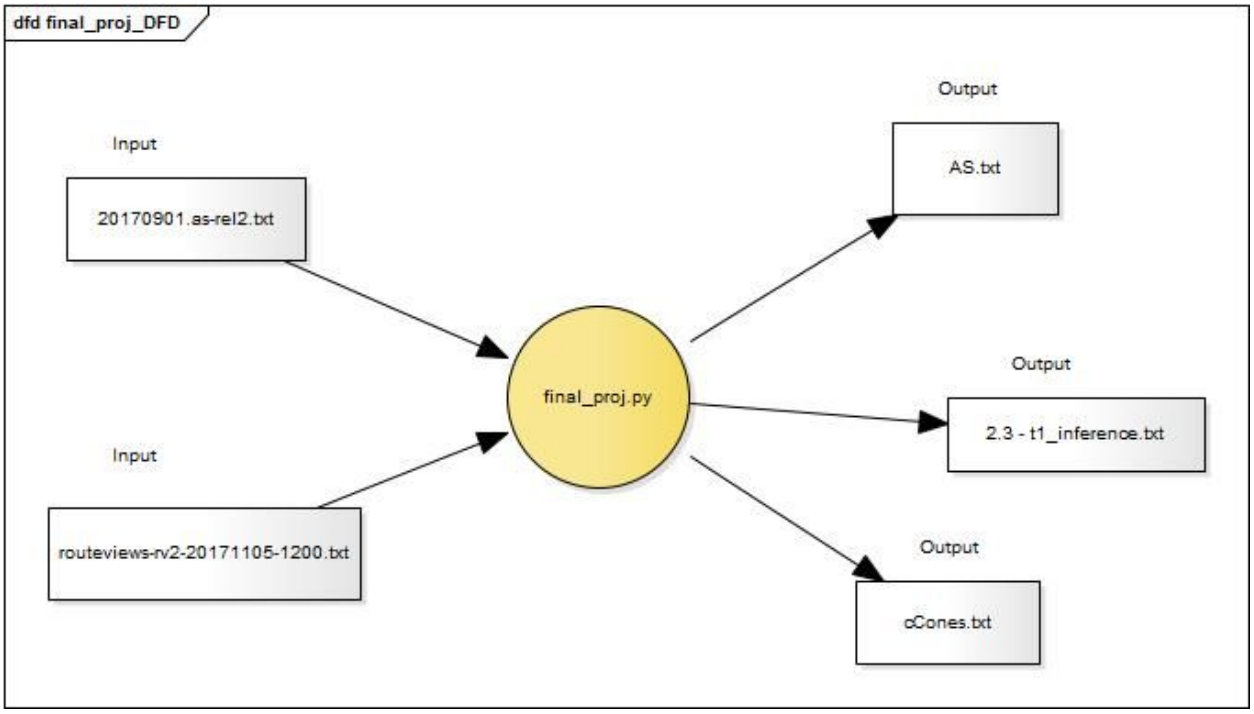
Thao Vo
Manuel Steele

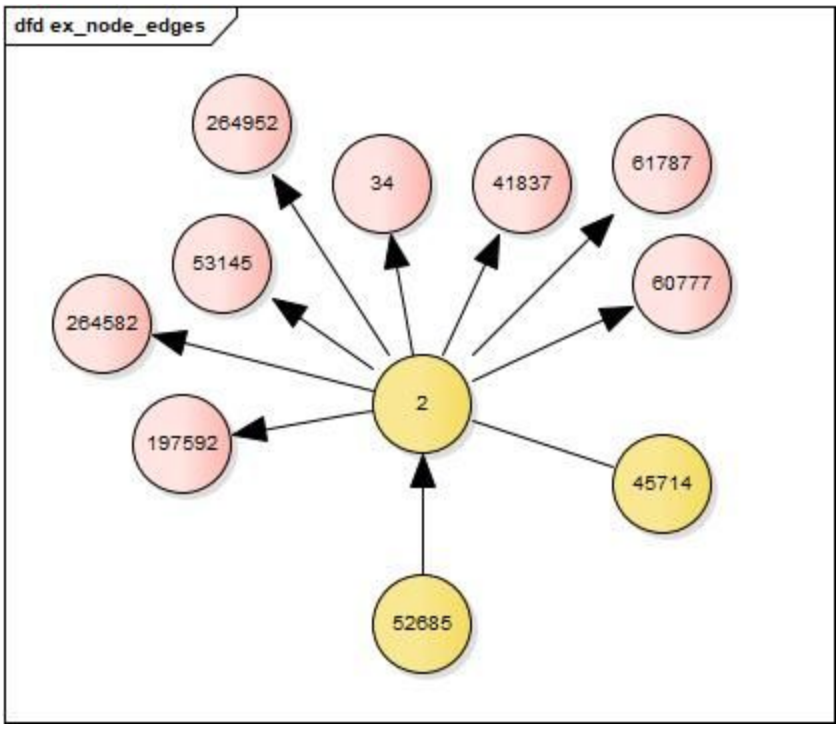
Fall 2017

1.0 Introduction

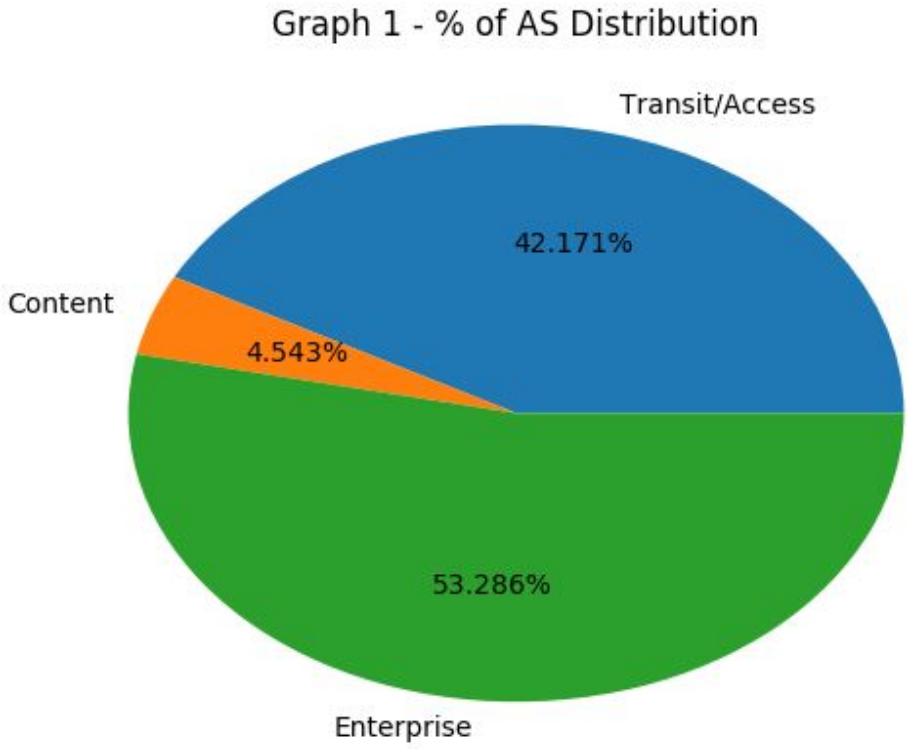
In this project, we analyzed data from the Center for Applied Internet Data Analysis (CAIDA) website. By using multiple methods of classification and analyzation, we were able to learn more about the distribution of Autonomous System and its different classification such as transit/access, content, and enterprise. The work was divided equally between both partners: Thao Vo and Manuel Steele. We were frequently on the phone to talking about the best method to solve the question, do unit testing, debugging, and essentially writing the report. Below are all of our findings.

1.1 Algorithm





2.1 Autonomous System (AS) Classification



Graph 1: % distribution of ASes to the 3 classes

	Transit/Access	Content	Enterprise	Total
Count	21,721	2,340	27,446	51,507
Percentage	42.171%	4.543%	53.285%	100%

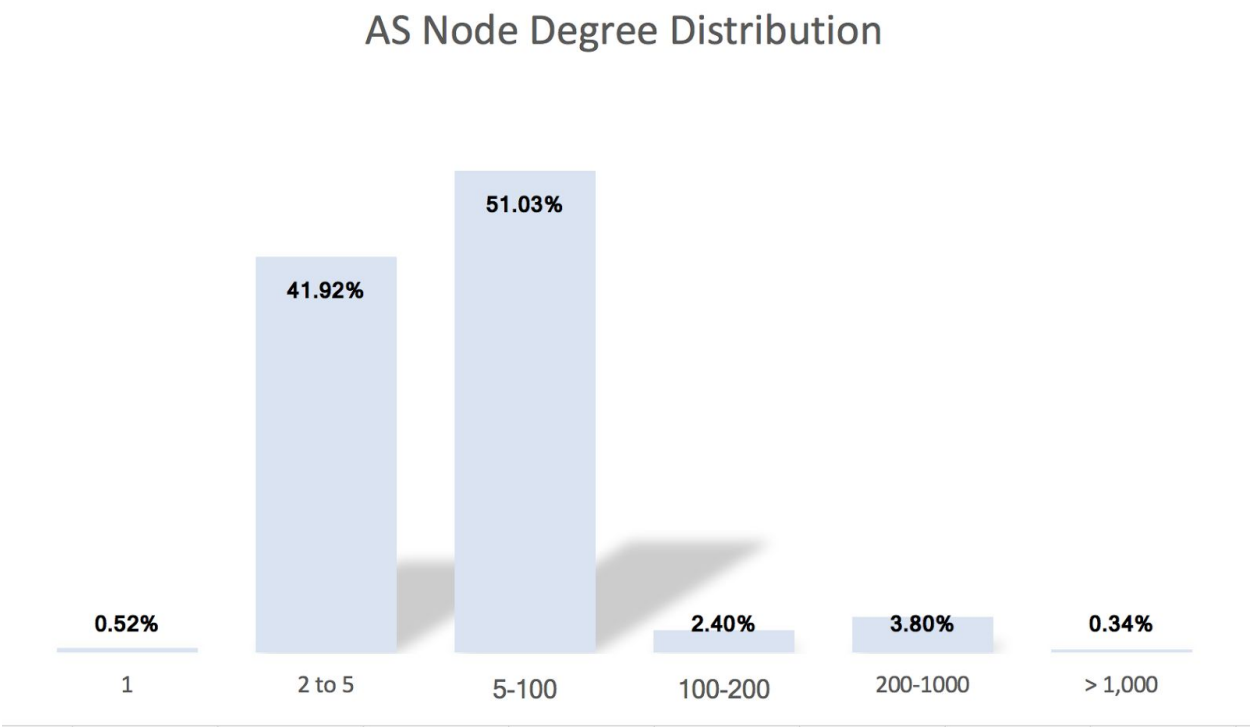
CAIDA data shows that the largest member of the ASes' space is at roughly 53.286%, the next is Transit at 42.171%, and the least is Content at 4.543%.

2.2 Topology Inference Through AS links

Graph 2

We took a lot of effort to develop the algorithm for this. We also manually did a unit test by picking a number of the list, and run it through manually to get the "supposed to be right" result. Initially, our code's calculation did not includes the provider. However, after a discussion question that Manuel posted on D2L, we were told to not exclude the pink nodes from the example below.

Therefore, we computed the connectivity degree irrespective of direction (provider, customer or peer). Ultimately, we are interested in the overall reachability of an AS by customers, providers, and peers and calculated them based on that. From then on, we retrieved:

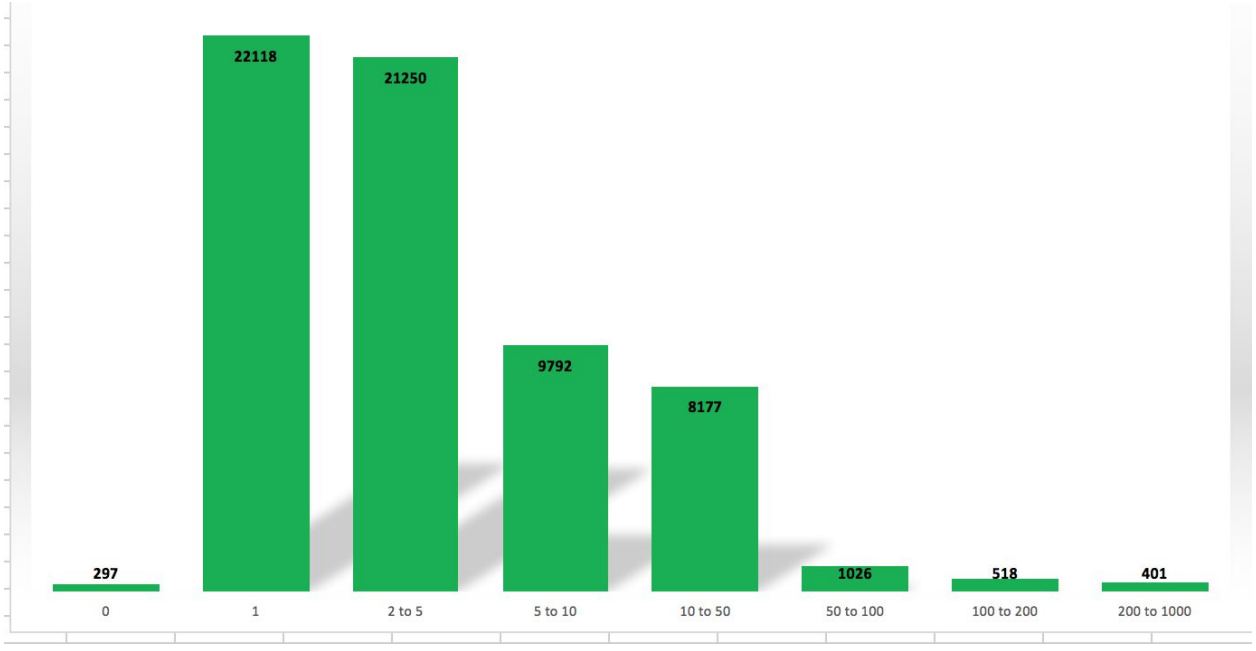


Graph 2: Histogram of AS node degree distribution using 1,2-5, 50-100, 100-200, 200-1000, <1000

	1	2-5	5-100	100-200	200-1000	>1000
Count	64	5,110	6,221	292	463	41

From this, we noticed that most AS are in the 5-100 degrees bin. That means that most ASes have about 2-5 connections with other nodes. From then, it becomes 5-100, 100-200, 200-1000, and >1000 and the lowest is just 1 single node. Most of the AS nodes in the 5-100 bins are mostly made up of 5. So if it was 6-100, the 2-5 bin would actually have the highest.

Graph 3



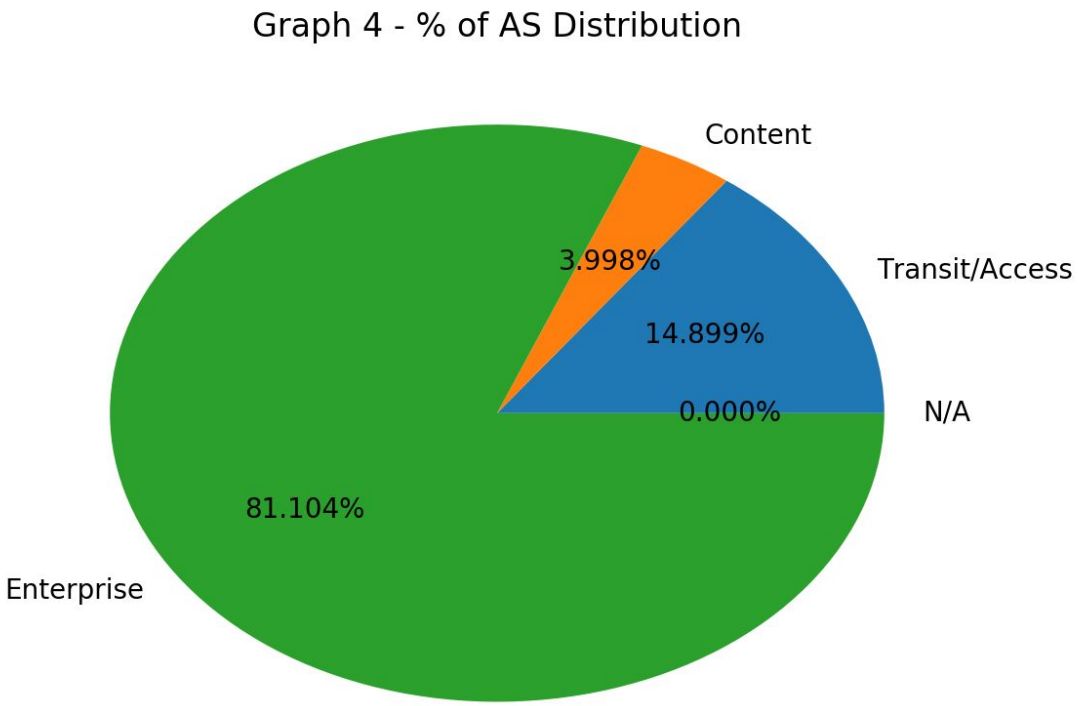
Graph 3 - Histogram of AS Node IP Prefixes Distribution

Observation:

We can see that most ASes falls within the #1 bin, which has 22,118 individual ASes. So we know that the most ASes are only connected to 1 other node that's either their peer, customer, or

provider. From that point on, more ASes are within the 2-5 prefixes bin. And it decreases as the number of prefixes goes up.

Graph 4



Graph 4: Recreate pie chart shows % distribution of ASes with criterias

Criterias:
Enterprise ASes: any AS with degree less than or equal to 2 and no customers or peers
Content AS: any AS with no customers and at least 1 peer
Transist: any AS with at least 1 customer

Observation:
From this graph, the data distribution seems to be consistent with Graph 1 where Enterprise remains as the largest. However, instead of 53.27%, it have jumped all the way to 81.104%. Transit data dropped down to around 14.90%, which I think we could say most of these transit classes was converted into enterprise. Content stayed pretty consistent with the lowest percentage.

2.3 Inference of T1 ASes

After running the algorithm, we only got 5. We used the AS # from the *routeviews-rv2-20171105-1200.txt* file and matched it along with the *20170401.as-org2info.txt* from the given website. The initial results are in black and results after adding the rest is in red after we found out that we should scan for more node after the first is disconnected.

Rank	**AS #**	**Organization**
1	6939	HURRICANE
2	174	COGENT-174
3	3356	LEVEL3
4	3549	LVLT-3549
5	7018	ATT-INTERNET4
6	209	CENTURYLINK-US-LEGACY-QWEST
7	2914	NTT-COMMUNICATIONS-2914
8	6461	ZAYO-6461
9	1299	TELIANET
10	3257	GTT-BACKBONE

Table 1: T1 list and the first 10 ASes that were added to S.

2.4 EC - Customer Cones and AS Rank

Table 2 - Ranked based on #'s of ASes

	AS INFO			Customer Cone					
				#'s			%'s		
Rank	**AS#**	**AS name**	**AS degree**	**#s ASes**	**IP Prefix**	**IPs**	**Percent ASes**	**IP Prefix**	**Ips**
1	174	COGENT-174	5366	50818	630799	2940723933	85.289%	86.503%	74.477%
2	3356	LEVEL3	4892	50620	632857	2982967491	84.957%	86.785%	75.547%
3	1299	TELIANET	1608	48814	627540	2974419396	81.926%	86.056%	75.331%
4	21320	GEANT_IAS_VRF	59	44556	576893	2686912808	74.780%	79.110%	68.049%
5	20965	GEANT	102	44555	576892	2686910760	74.778%	79.110%	68.049%
6	2914	NTT-COMMUNICATIONS-2914		43068	573485	2599834860	72.282%	78.643%	65.844%
7	6453	AS6453	719	42132	579207	2735813491	70.711%	79.428%	69.288%
8	3257	GTT-BACKBONE	1566	42120	566619	2433081047	70.691%	77.701%	61.621%
9	5511	Opentransit	163	40868	553878	2529540574	68.590%	75.954%	64.063%
10	3491	BTN-ASN	602	39810	548313	2480087456	66.814%	75.191%	62.811%
11	6939	HURRICANE	6512	39196	540804	2429643974	65.784%	74.161%	61.533%
12	209	CENTURYLINK-US-LEGACY-QWEST	1892	37858	516392	2251994038	63.538%	70.814%	57.034%
13	2603	NORDUNET	712	37846	502339	2147195779	63.518%	68.887%	54.380%

14	701	UUNET	1234	37786	520964	2454016293	63.417%	71.441%	62.151%
15	6762	SEABONE-NET	412	36560	506263	2156897623	61.360%	69.425%	54.626%

Table 2: Top 15 ASes ranked by customer cone in #'s of ASes that they can reach using p2c links

Table 2 - Ranked based on% of IPs

AS INFO					**Customer Cone**				
					#'s			**%'s**	
AS Rank	**AS#**	**AS name**	**AS degree**	**#s ASes**	**IP Prefix**	**IPs**	**Percent ASes**	**IP Prefix**	**Ips**
1	3356	LEVEL3	4892	50620	632857	2982967491	84.957%	86.785%	75.547%
2	1299	TELIANET	1608	48814	627540	2974419396	81.926%	86.056%	75.331%
3	174	COGENT-174	5366	50818	630799	2940723933	85.289%	86.503%	74.477%
4	6453	AS6453	719	42132	579207	2735813491	70.711%	79.428%	69.288%
5	21320	GEANT_IAS_VRF	59	44556	576893	2686912808	74.780%	79.110%	68.049%
6	20965	GEANT	102	44555	576892	2686910760	74.778%	79.110%	68.049%
7	2914	NTT-COMMUNICATIONS-2914	1649	43068	573485	2599834860	72.282%	78.643%	65.844%
8	5511	Opentransit	163	40868	553878	2529540574	68.590%	75.954%	64.063%
9	3491	BTN-ASN	602	39810	548313	2480087456	66.814%	75.191%	62.811%
10	701	UUNET	1234	37786	520964	2454016293	63.417%	71.441%	62.151%
11	3257	GTT-BACKBONE	1566	42120	566619	2433081047	70.691%	77.701%	61.621%
12	6939	HURRICANE	6512	39196	540804	2429643974	65.784%	74.161%	61.533%
13	1239	SPRINTLINK	451	35984	519689	2384943366	60.393%	71.266%	60.401%

14	209	CENTURYLINK-US-LEGACY-QWEST	1892	37858	516392	2251994038	63.538%	70.814%	57.034%
15	4766	KIXS-AS-KR	524	34808	492577	2235850074	58.419%	67.548%	56.625%

Most of the organization are the same in both tables 2 & 3 from above. Their ranking varied.