1 | 2 | 3
10 | 10 | 10

6.829 **Dan Ports**

Computer Networks drkp@mit.edu

2005/09/27 Collaborators: {**amdragon**}

Problem Set 2

Problem 1:

1) Consider a full mesh. P1 is satisfied: every router advertises to every other router the best routes it sees from eBGP, so every router sees the best routes from each eBGP router, and hence can choose the best route. P2 is satisfied: we assume that the paths learned via eBGP are loop-free. The IGP shortest-paths are necessarily also loop-free. Since iBGP does not readvertise iBGP-learned paths to other routers, every router's path to any destination will consist of an IGP path to a route learned by eBGP, both components of which are loop-free.

2) This configuration violates P1 and P2. C1 learns a route to d via R1, and C2 learns a route of the same cost to d via R2. However, C1's shortest path to R1 is via C2 (cost 3 rather than 5), and similarly C2's shortest path to R2 is via C1. So if C1 tries to send a packet to d, it will send it to C2. If C1 had complete knowledge of eBGP routes, it would choose to send it to R2 to forward to d (which has cost 2 rather than 3). So P1 is violated. Moreover, when it forwards the packet to C2, C2 will forward it back to C1, creating a routing loop, since its best route to d is via C1 and R2. So P2 is violated.

3) In this case, the reflectors will forward routes to d via both R1 and R2 to their clients. So R1 and R2 will choose the direct routes, C1 will choose the optimal route of forwarding to R2, and C2 will forward to R1. This is optimal and loop-free, so P1 and P2 are satisfied.

4) If the iBGP configuration satisfies this property, then every router A will know about the best routes from every egress router B: in cases (a), (b), and (c), A learns the route directly from B, and in case (d), B sends the route to the reflector which advertises it to A. So A will learn the best routes from every egress router B, and choose the best of these, satisfying P1. P2is satisfied because A will forward the packet along the shortest path to some egress router B. Because the iBGP configuration satisfies this property, the next hop router will also know the best route, and so the distance to the egress router is always decreasing. This progress condition ensures that forwarding loops are not possible, satisfying P2.

Problem 2:

1) *a)* MIT is AS #3.

b) The best next hop is `4.68.0.243`; the router knows how to reach this IP via an IGP-discovered route.

c) A packet needs to travel through one AS (AS #3356) to reach MIT.

d) A traceroute from MIT to route-views follows:

```
traceroute to route-views.oregon-ix.net (128.223.60.103), 64 hops max, 40 byte packets
 1  wireless.kalgan.csail.mit.edu (128.30.4.2)  3.843 ms  2.533 ms  2.112 ms
 2  kalgan.trantor.csail.mit.edu (128.30.0.245)  3.791 ms  1.507 ms  9.202 ms
 3  b24-rtr-2-csail.mit.edu (18.4.7.1)  1.599 ms  3.370 ms  1.668 ms
 4  external-rtr-1-backbone.mit.edu (18.168.0.18)  13.945 ms  5.727 ms  1.914 ms
 5  external-rtr-2-backbone.mit.edu (18.168.0.27)  1.776 ms  2.975 ms  8.062 ms
```

```
 6  nox230gw1-vl-526-nox-mit.nox.org (192.5.89.89)  3.644 ms  10.391 ms  9.200 ms
 7  nox230gw1-peer-nox-nox-192-5-89-10.nox.org (192.5.89.10)  7.672 ms  32.683 ms  12.576 ms
 8  chinng-nycmng.abilene.ucaid.edu (198.32.8.82)  42.102 ms  29.618 ms  27.834 ms
 9  iplsng-chinng.abilene.ucaid.edu (198.32.8.77)  235.201 ms  251.583 ms  238.747 ms
10  kscyng-iplsng.abilene.ucaid.edu (198.32.8.81)  44.770 ms  44.764 ms  42.351 ms
11  dnvrng-kscyng.abilene.ucaid.edu (198.32.8.13)  65.430 ms  101.479 ms  71.056 ms
12  snvang-dnvrng.abilene.ucaid.edu (198.32.8.1)  78.085 ms  81.139 ms  78.952 ms
13  pos-1-0.core0.eug.oregon-gigapop.net (198.32.163.17)  88.204 ms  89.724 ms  89.129 ms
14  uo-0.eug.oregon-gigapop.net (198.32.163.147)  128.623 ms  120.511 ms  123.037 ms
15  ge-5-1.uonet1-gw.uoregon.edu (128.223.2.1)  102.757 ms ge-5-1.uonet2-gw.uoregon.edu (128.223
16  g0-1.route-views.routeviews.org (128.223.60.103)  101.156 ms *  88.975 ms
```

This route passes through MIT (AS #3), NOX/GIGAPOP-NE (AS #10578), Abilene (AS #11537), UO (AS #4600), and UONET (AS #3582). This is not the same as the reverse of the route in the trace data.

e) The AS path from MIT to routeviews is not the same as the reverse of the path from routeviews to MIT because MIT prefers a different route (via Abilene), presumably because of a `LOCAL PREF` option in MIT's router.

The traceroute actually matches the AS path: the `nox.org` addresses are part of AS #10578 and the Abilene hosts are AS #11537. The AS appears as `null` in the traceroute because the `whois` database consulted by the traceroute program did not include these ASes, but searching a more comprehensive database revealed the AS numbers. The brief excursion through AS #40 at the beginning can be explained because AS #40 is also part of MIT, so this is still an internal hop.

f) There are 51 routes from routeviews to MIT.

g) The router prefers a route via Level3 (AS #3356) to reach MIT, because this route has the shortest path length (two ASes) and lowest metric (zero).

h) MIT's neighboring ASes are:

1. AS #174 (`COGENT-PSI-1`) — Cogent Communications
2. AS #1239 (`SPRINTLINK`) — Sprint
3. AS #3356 (`LEVEL3`) — Level 3 Communications
4. AS #10578 (`GIGAPOP-NE`) — Harvard University

2) *a)* This can be checked by testing whether

$$A_i >> (32 - m) == A >> (32 - m)$$

b) The first such address is `192.0.32.0/21`. This corresponds to $2^3 = 8$ class C networks. So CIDR is reducing at most 8 routing table entries to 1 here, giving a maximum savings of 7. However, some of the address space in the /21 block might be unused, so some of the class C blocks might not need routing table entries, in which case the savings would be less.

c)

1. The prefixes `4.21.104.0/24` and `4.21.112.0/23` both have routes via the AS path 16150 → 1239 → 7018 → 26207. This seems to be because the same organization (`AQUILENT`) owns both netblocks, but was not allocated a single contiguous block; the blocks in between, such as `4.21.106.0/23` are owned by different organizations.
2. The contiguous blocks `8.10.114.0/24` and `8.10.115.0/24` both have routes via the same AS path 3356 → 21904. This deaggregation might be useful in order to allow the possibility of advertising different routes to each of the /24 blocks, for load balancing or similar reasons. But it increases the size of every router's routing table.

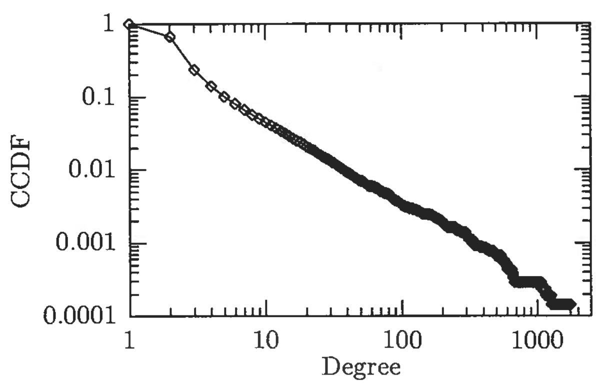
3) *a)* Having only the destination network and mask would not provide much useful information, but it would be possible to determine which IP addresses are unroutable and hence make some guesses about the number of hosts on the Internet. Since CIDR supports aggregation, it's not entirely unreasonable to use the number of destination network masks in the routing table as an estimate of the number of distinct networks, but this would not be especially accurate.

b) This would include the best next-hop and AS path for every routable address, so it would allow one to determine the number of ASes, and some of the links between ASes. It would be possible to determine how packets were routed from the Oregon Exchange at the time of each snapshot. It would be possible to attempt to infer AS relationships from the AS paths as in Problem 3, though this provides only a partial view since all the other paths are missing.

c) This provides all the information that can be inferred from (b), except for information about the next hop IP address, and metric/weight/MED/local-pref information. It also provides a view of AS links, so could be used to infer AS relationships as in Problem 3, using either the degree ranking method or the graph pruning method (though only with a single vantage point).

Problem 3:

1) A CCDF of AS degree follows:



Top 10 ASes by degree

Degree	AS #	Name
2404	701	ALTERNET-AS / UUNET Technologies, Inc
2001	7018	ATT-INTERNET4 / AT&T WorldNet Services
1763	1239	SPRINTLINK / Sprint
1267	3356	LEVEL3 / Level 3 Communications, LLC
1148	209	ASN-QWEST / Qwest
1063	174	COGENT-PSI-1 / Cogent Communications
675	3549	GBLX / Global Crossing
659	6461	ABOVENET / Abovenet Communications, Inc
646	4323	TWTC / Time Warner Telecom
598	7132	SBIS-AS / SBC Internet Services

2) Let "A obtains transit from B" be denoted as $A \rightarrow B$; "A provides transit to B" be denoted as $A \leftarrow B$; and a sibling relationship between A and B as $A \leftrightarrow B$.
a) $16150 \rightarrow 1239 \rightarrow 701 \leftarrow 703 \leftarrow 80$

b) $7660 \longrightarrow 2516 \longrightarrow 1239 \longrightarrow 7018 \longleftarrow 2386$

c) $3277 \leftrightarrow 3267 \leftrightarrow 3343 \longrightarrow 1299 \longrightarrow 3549 \longleftarrow 206$

d) $4513 \leftrightarrow 7911 \longleftarrow 3320 \longleftarrow 8551$

3) *a)* The advantage of this ranking scheme is that it can identify provider-customer relationships even when the degree is not proportional to the AS's size; the paper gives the examples of a small ISP that purchases transit from many providers, and a large ISP that has peering relationships with only a few providers. However, it requires multiple vantage points to be effective.

b) The scheme requires multiple vantage points because some links may not be visible from every vantage point. For example, an ISP will not advertise transit routes it has purchased to its peers or other providers, so they will not see these routes. Also, multiple vantage points are necessary to identify peering relationships: a single vantage point will suggest that one peer is a customer of the other, and a different one will suggest the opposite.