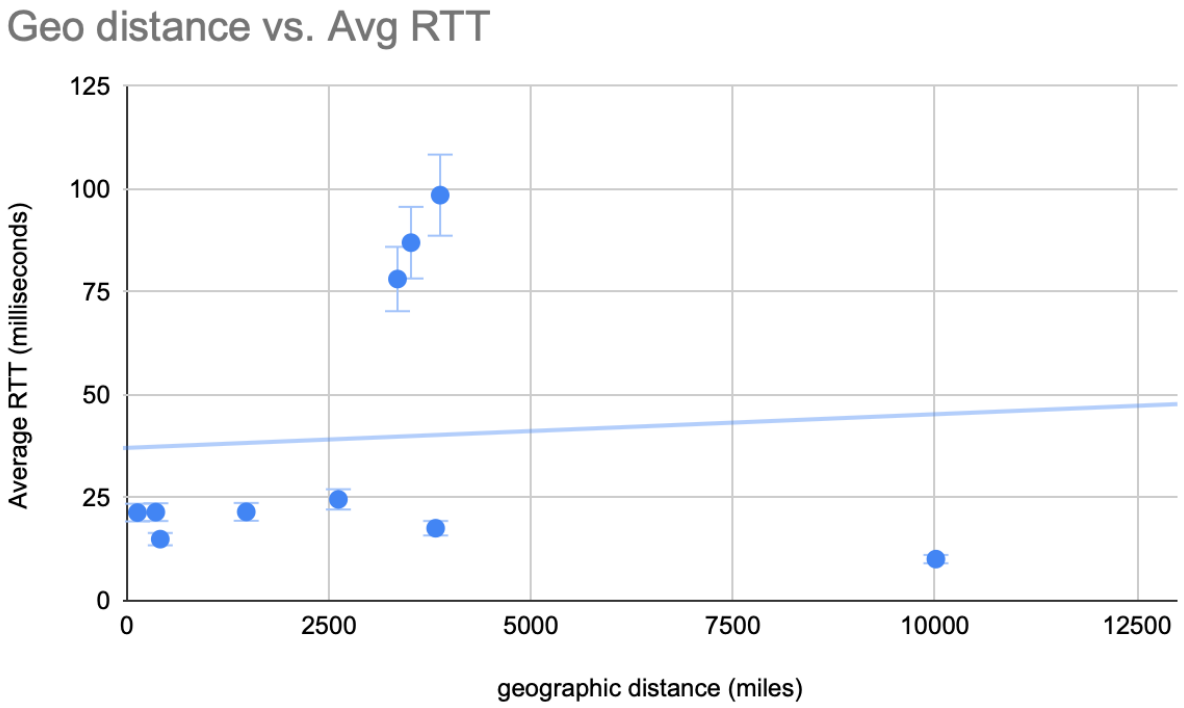
Project 4 Ping Feiran Chang & Nova Zhang

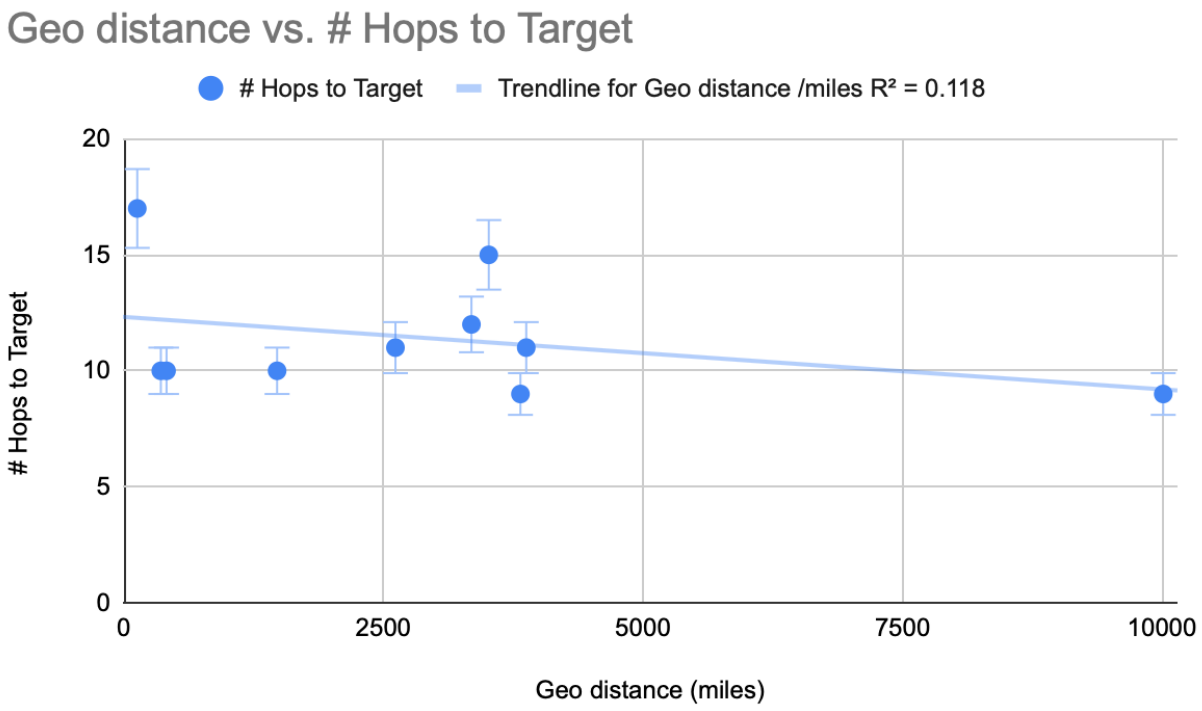


1. Are RTT and geographic distance correlated positively, negatively, or not at all? If applicable, also comment on the strength of the correlation (weak vs. strong).

   There's a weak positive correlation between RTT and geographic distance. As the geographic distance increases, the RTT tends to increase as well. Although we can see an outlier that has a 10010-mile geographic distance but only 10 ms RTT, it is reasonable to say that according to the performances of the other 9 IP addresses, there is a positive correlation between RTT and geographic distance.

2. Why do you think you observe this trend (or lack thereof)?

   We had a bunch of outliers. The average RTT of IP from California(PA), Zurich, and Sydney are shorter than others, which according to the trend above, we expected the average RTT to be longer. Another proof is that the R^2 value of the linear regression line is small, which indicates a weak linear correlation. Another possible reason is that the RTT time is highly correlated with the hops. It is possible that although the geographical distance is far, the hops the data used to travel provide higher transmission speed due to less congestive traffic or special transmission media like optic fiber.



1. Are # hops and geographic distance correlated positively, negatively, or not at all? If applicable, also comment on the strength of the correlation (weak vs. strong).

   There's a negative correlation between # hops and geographic distance. As the geographic distance increases, the number of hops tends to decrease. The correlation is very weak.

2. Why do you think you observe this trend (or lack thereof)?

   We did not get a clear linear correlation. The # hops of the last three IPs (Zurich, Falkenstein, Sydney) are lower than we expected based on the trend above. And the # hops of the IP at North Bergen (NJ) is much greater than we expected. Another proof is that the R^2 value of the linear regression line is small, which indicates a weak linear correlation. And the slope of that linear regression line is negative which indicates a lack of proof of this trend. One guess to explain this is that the locations near us are where population density is high and a lot of hops are gathered. Therefore, the data passes many hops but only takes a relatively small amount of time. On the contrary, locations that are far away from us may use submarine cables, which only involve one or two hops, but take more time.

Supporting information:

Ping data

	Domain/IP	Geo distance /miles	Repeat 1	Repeat 2	Repeat 3	Avg RTT	Avg RTT in miliec	Geo Location
1	161.35.252.68	127	0.02416300774	0.02411603928	0.01595115662	0.02141006788	21.41006788	North Bergen, NJ
2	172.67.195.193	354	0.02462172508	0.01621794701	0.02371716499	0.02151894569	21.51894569	Toronto, Canada
3	185.199.110.153	409	0.01831293106	0.01057100296	0.01598381996	0.01495591799	14.95591799	California, PA
4	209.244.0.3	1474	0.02946019173	0.01652193069	0.01877236366	0.02158482869	21.58482869	Dallas, TA
5	8.8.8.8	2614	0.02847886086	0.02733802795	0.01803326607	0.02461671829	24.61671829	Mountain View, CA
6	146.75.74.2	3346	0.07534122467	0.08229780197	0.07684330233	0.07816068331	78.16068331	London, Engalnd
7	18.155.129.64	3513	0.08410429955	0.09279608727	0.08409690857	0.08699909846	86.99909846	Paris, France
8	9.9.9.9	3818	0.01847982407	0.01394319534	0.02039885521	0.01760729154	17.60729154	Zurich, Switzerland
9	148.251.136.139	3874	0.09971785545	0.09759402275	0.09831786156	0.09854324659	98.54324659	Falkenstein, Germany
10	1.0.0.1	10010	0.009139299393	0.008632183075	0.01255583763	0.0101091067	10.1091067	Sydney, Australia

Traceroute data

	Domain/IP	Geo distance /miles	# Hops to Target
1	161.35.252.68	127	17
2	172.67.195.193	354	10
3	185.199.110.153	409	10
4	209.244.0.3	1474	10
5	8.8.8.8	2614	11
6	146.75.74.2	3346	12
7	18.155.129.64	3513	15
8	9.9.9.9	3818	9
9	148.251.136.139	3874	11
10	1.0.0.1	10010	9