https://www.overleaf.com/project/634049e98eaeda173074c9cb

DEPARTAMENTO DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Practico II: *Rutas en Internet*

19 de octubre de 2022 — Teoría de las Comunicaciones

Integrante	LU	Correo electrónico
Monteys, Lautaro	[redacted]	[redacted]
Vieytes, Vicente	[redacted]	[redacted]
Oca, Mariano	[redacted]	[redacted]
Chami, Uriel Alberto	[redacted]	[redacted]

Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria - (Pabellón I)
Intendente Güiraldes 2610 - C1428EGA
Ciudad Autónoma de Buenos Aires - Rep. Argentina
Tel/Fax: (++54 +11) 4576-3300
http://www.exactas.uba.ar

Índice

1. Introducción **2**

2. Métodos y condiciones de los experimentos **3**

3. Resultados y análisis de los experimentos **4**

4. Conclusión **7**

5. Anexo **9**

1. Introducción

En este informe analizaremos la topología y los protocolos de nivel red. Para ello utilizaremos el método conocido como *traceroute*, específicamente la implementación con *echo request - echo reply* el cual consiste en enviar múltiples paquetes ICMP de *echo request* con TTL cada vez mayor. Cada paquete recibirá en respuesta (idealmente) un mensaje de TTL timed out como el protocolo ICMP establece. En una red "lineal" como la de la figura 1 es claro entender por qué y cómo este método en términos generales funciona. Si nuestros paquetes viajaran de esta manera lineal, tendríamos como resultado: 4 envíos de paquetes ICMP, con 4 respuestas del tipo TTL expired, que además traen consigo un IP fuente que nos dice quién es ese router respondiendo. Además tomando el tiempo entre envío y respuesta, tendremos el Round Trip Time (RTT).

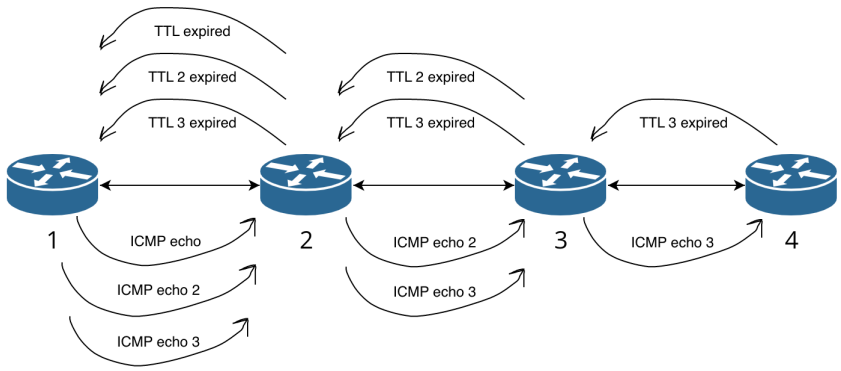

Figura 1: Red ideal lineal

Sin embargo, este caso ideal puede no darse a causa de varias posibles anomalías de traceroute:

- Routers faltantes: algunos routers pueden no contestar cuando expire un TTL o estar protegidos por un firewall lo que resulta en nodos faltantes en el traceroute.
- Destino faltante: Análogamente a la anomalía anterior. En este caso genera que el traceroute no se detenga hasta no alcanzar el TTL máximo.
- RTTs incorrectos: se puede dar cuando la ruta de ida no es la misma que de vuelta.
- Conexiones inexistentes: cuando algunos paquetes viajan por una ruta y otros por otra. Se puede dar por balance de carga.
- Routers repetidos: en el caso de que el balance de carga se utilice para caminos de longitudes distintas.

Por otro lado estudiaremos los tiempos de respuesta entre un paso y el otro buscando outliers para tratar de encontrar los saltos intercontinentales en las redes. Lo haremos utilizando el método Thompson tau modificado (método Cimbala [4]), esta nos indicará qué tiempos son outliers. Mirando estos valores, utilizando servicios que asocian IPs con locaciones físicas y también teniendo en cuenta el cableado internacional conocido buscaremos los saltos entre un continente y el otro.

2. Métodos y condiciones de los experimentos

Para la ejecución de los experimentos implementamos traceroute con paquetes ICMP haciendo uso de la biblioteca Scapy. Nuestro código ejecuta un camino completo de traceroute aumentando el TTL de a 1 llegando como máximo a 29, esto se repite 30 veces para finalmente tomar el promedio de los RTTs obtenidos. Las IP asociadas a cada TTL es la IP con más apariciones para ese mensaje. Sin embargo no nos encontramos con ejemplos donde más de un router responda a cierto TTL, lo que no necesariamente significa que nuestros paquetes recorran un solo camino si no que sólo pudimos observar un camino dado que no todos los routers responden. El código utilizado se encuentra en la sección de referencias [1].

Nuestras muestras fueron tomadas el 15/10/2022 por la noche. Tomando los siguientes destinos:

- *Australian College of Physical Education*: `acpe.edu.au`
- *Indian Institute of Technology*: `home.iitd.ac.in`
- *University of South Africa*: `unisa.ac.za`
- *Qatar University*: `qu.edu.qa`

3. Resultados y análisis de los experimentos

Las siguientes tablas muestran los resultados de tomar la IP del router con más apariciones para cada TTL y calcular el RTT promedio para cada uno de estos.

TTL	Dirección IP	RTT ms	Outlier	País
1	192.168.1.1	18.3		Arg
2	181.96.112.126	16.46		Arg
3	181.88.170.178	17.72		Arg
6	181.96.103.168	17.82		Arg
7	195.22.220.56	38.32	X	Arg
8	89.221.41.171	144.44	X	EEUU
10	154.54.88.233	154.11		Brasil
11	154.54.84.1	181.95	X	EEUU
12	154.54.30.162	190.58		EEUU
13	154.54.42.65	189.42		EEUU
14	154.54.44.86	191.08		EEUU
15	154.54.31.190	203.12		EEUU
16	154.54.1.162	201.44		EEUU
17	38.142.245.26	201.97		EEUU
20	27.122.113.133	396.47	X	Australia
21	103.252.152.253	396.99		Australia
22	103.252.152.183	495.08	X	Australia
23	103.252.152.190	481.26		Australia
24	110.232.143.40	487.72		Australia

Cuadro 1: Resultados promediados del traceroute a un IP de Australia.

TTL	Dirección IP	RTT ms	Outlier	País
1	192.168.1.1	22.02		Arg
2	181.96.112.126	16.68		Arg
3	181.88.170.212	15.48		Arg
6	181.96.113.234	309.1	X	Arg
7	195.22.220.56	34.24		Arg
9	149.3.181.65	64.12		Italia
10	129.250.5.24	46.57		EEUU
11	129.250.2.12	151.51	X	EEUU
12	129.250.6.81	166.57		EEUU
13	128.241.7.159	167.03		EEUU
14	103.198.140.55	390.71	X	Singapur
15	103.198.140.40	391.07		Singapur
16	103.198.140.55	385.7		Singapur
17	49.45.4.252	401.3		India

Cuadro 2: Resultados promediados del traceroute a un IP de India

TTL	Dirección IP	RTT [ms]	Outlier	País
1	192.168.1.1	14.18		Arg
2	181.96.112.126	16.75		Arg
3	181.88.171.148	14.7		Arg
6	181.96.113.234	42.84		Arg
7	195.22.220.56	20.86		Arg
8	195.22.219.65	45.2		Brasil
9	149.3.181.65	46.04		Italia
10	129.250.2.196	153.35	X	EEUU
11	129.250.2.108	152.68		EEUU
12	129.250.200.114	161.61		EEUU
13	170.39.8.30	154.77		EEUU
15	155.232.1.149	455.39	X	Sudáfrica
16	155.232.1.97	425.32		Sudáfrica
17	192.96.2.249	409.79		Sudáfrica
18	163.200.81.55	410.56		Sudáfrica

Cuadro 3: Resultados promedio del traceroute a un IP de Sudáfrica.

TTL	Dirección IP	RTT [ms]	Outlier	País
1	192.168.1.1	15.88		Arg
2	181.96.112.126	15.95		Arg
3	181.88.171.148	14.28		Arg
6	181.96.103.168	18.91		Arg
7	195.22.220.56	19.05		Arg
8	213.144.170.29	269.26	X	Italia
9	213.144.170.111	277.19		Italia
10	89.211.3.25	372.32	X	Qatar
11	89.211.0.65	374.52		Qatar
12	89.211.1.238	399.33	X	Qatar

Cuadro 4: Resultados promedio del traceroute a un IP de Qatar

En todos los casos se observa un return time total de aproximadamente 400 milisegundos, si bien la cantidad de saltos varía significativamente entre rutas. Se encuentran marcados las IPs para las cuales el salto para llegar a ellas desde la anterior registró un RTT diferencial que resultó ser un outlier según el método Cimbala.

Una anomalía que se observó en todos los experimentos es la de "missing hops", esto es cuando falta la respuesta de `Time Exceeded` para algún TTL intermedio, por ejemplo en el traceroute a Sudáfrica faltó la respuesta para los TTLs 3, 4, 5 y 14. Esto se debe probablemente a firewalls que bloquean estas respuestas ICMP o a routers MPLS. Aparte de esto no se observó comportamiento anómalo a simple vista mientras corriamos el programa, ya que el RTT fue creciente en todos los experimentos y no se observaron loops ni RTTs excesivamente sospechosos. Esto no significa que no haya más anomalías ya que algunas son difíciles de detectar. [1]

[1] Martin Erich Jobst, *Traceroute Anomalies* (2012)

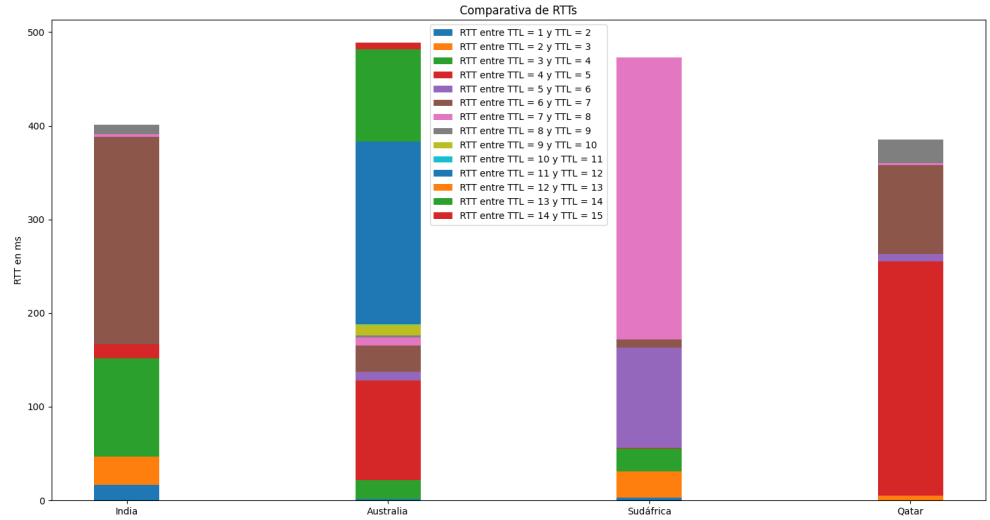

Figura 2: RTTs para cada ruta

En la Figura 2 se representan comparativamente los valores de RTT de cada mensaje de la ruta. Una buena explicación para los repentinos incrementos en RTT que se observan en el gráfico y en las tablas es la de los saltos oceánicos que realizan los paquetes en el ruteo.

Nos interesa estimar qué saltos en RTT son suficientemente significativos como para indicar un posible salto oceánico (el cual debería destacar sobre cualquier tramo continental). Para determinar estos, aplicamos el método Thompson tau modificado (Cimbala) [2] al valor de cada salto entre los RTT y obtuvimos los outliers de cada muestra experimental.

¿La distribución de RTT entre saltos presenta outliers según el método de Cimbala? ¿Cuántos?

Sí, los saltos de RTT entre las siguientes direcciones IP fueron considerados outliers por este método:

- IP de Australia, 5 outliers: entre 181.96.103.168 y 195.22.220.56, entre 195.22.220.56 y 89.221.41.171, entre 154.54.88.233 y 154.54.84.1, entre 38.142.245.26 y 27.122.113.133, y entre 103.252.152.253 y 103.252.152.183.
- IP de India, 3 outliers: entre 181.88.170.212 y 181.96.113.234, entre 129.250.5.24 y 129.250.2.12, y entre 128.241.7.159 y 103.198.140.55
- IP de Sudáfrica, 2 outliers: entre 149.3.181.65 y 129.250.2.196, y entre 170.39.8.30 y 155.232.1.149.
- IP de Qatar, 3 outliers: entre 195.22.220.56 y 213.144.170.29, 213.144.170.111 y 89.211.3.25, y entre 89.211.0.65 y 89.211.1.238.

Por otro lado utilizamos varias herramientas en linea de geolocalización por IP [2] para determinar la posición de cada dirección, y la biblioteca Plotly para producir el recorrido aproximado de cada ráfaga de paquetes.

[2] John M. Cimbala, *Outliers* (2011)

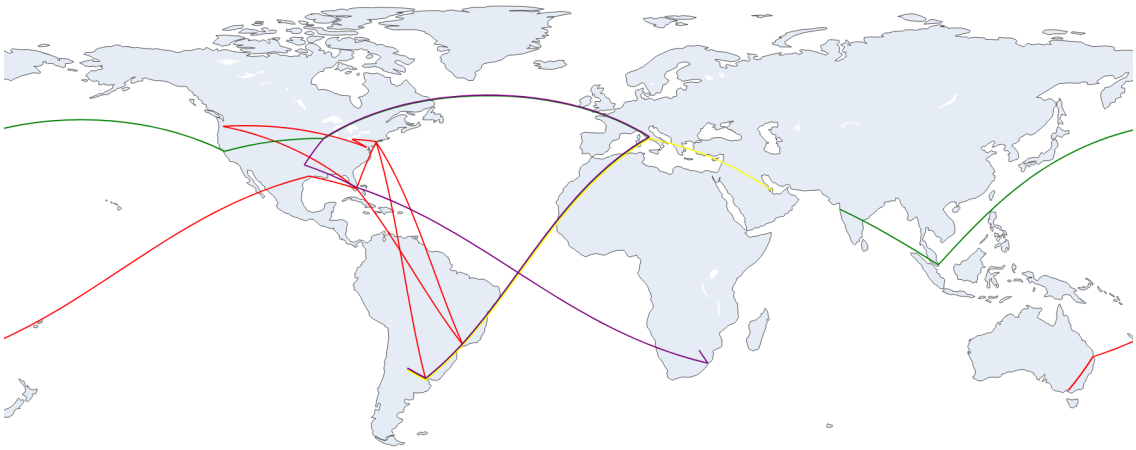

Figura 3: Mapa con las rutas aproximadas de cada paquete para llegar al ip destino. Servimos una versión interactiva de este mapa en https://traceroute-map.000webhostapp.com/ donde se puede hacer zoom y aislar cada ruta haciendo doble click en una IP. Recomendamos fuertemente esta manera de inspeccionar el mapa.

¿Se corresponden los outliers con los enlaces intercontinentales? ¿Cuántos falsos positivos y falsos negativos hay?

Comparando los resultados del estimado estadístico de outliers con la información obtenida con geolocalización de IPs encontramos las siguientes diferencias:

- IP de Australia: 5 outliers, 1 salto interoceánico (de Estados Unidos a Australia), 4 falsos positivos. La gran cantidad de falsos positivos en este caso se debe a que en la traza hay muchos reportes de RTTs extremadamente bajos, lo que genera que RTTs de 20ms en adelante sean marcados como outliers.

- IP de India: 3 outliers, 3 saltos interoceánicos (de Brasil a Italia, de Italia a Estados Unidos y de Estados Unidos a Singapur), 2 falsos positivos, 1 falso negativo. El método sólo detecta el salto de Estados Unidos a Singapur. El RTT entre dos IPs de Argentina (3 y 6) es anormalmente mayor con 309 milisegundos, esta anomalía genera que se marque como outlier cuando no es un salto interoceánico. Se explica en parte porque en ese salto se incluyen otros dos que no son parte de la traza debido a que no respondieron. Pero aún así, es tanta la diferencia que otra causa como congestión se debe haber dado.

- IP de Sudáfrica: 2 outliers, 3 saltos interoceánicos (de Brasil a Italia, de Italia a Estados Unidos, y de Estados Unidos a Sudáfrica), 1 falso negativo. El método no reconoce el salto de Brasil a Italia ya que una anomalía causó un falso RTT por la cual la diferencia entre estas dos IPs resultó incorrectamente pequeña. De todas formas, hay que mencionar que los tres saltos pueden no haberse dado realmente y que el resultado haya sido consecuencia de una anomalía de *conexiones inexistentes*.

- IP de Qatar: 3 outliers, 1 salto interoceánico (sobre el Atlántico), 2 falsos positivos. El método reporta el salto correctamente pero, también, reporta dos falsos positivos entre una IP de Italia y otra de Qatar; además de entre dos IPs de Qatar. El primer falso positivo se debe a simple distancia y es un error esperable. Sin embargo, el segundo se debe a una latencia ligeramente superior entre dos routers del mismo país.

En total, la cantidad de falsos positivos utilizando este método fue de 8 y la de falsos negativos en principio fue 2. Es decir, reportó como tales el 71 % de los saltos interoceánicos presentes, pero

con alto número de falsos positivos. Sin embargo, ambos falsos negativos son por no reconocer un salto desde una supuesta IP de América del Sur a una supuesta IP de Italia, pero mirando los RTT de estas IP, como ambos son menores a 100ms, concluimos que se debe a un error de geolocalización y en realidad corresponden a un router de América del Sur. Por lo tanto, no son realmente saltos interoceánicos y la cantidad de falsos negativos es 0.

Esto se debe a la considerable cantidad de anomalías que se presentan en los traceroutes. El método no presenta una gran robustez frente a ellas y es propenso a errores, si bien es una buena primera aproximación al análisis. Además, no tiene forma de diferenciar entre una alta latencia debido a un salto interoceánico de una que se debe a que simplemente los nodos se encuentran a gran distancia, sin un océano que los separe.

¿Es posible mejorar las predicciones usando un valor de corte fijo para el valor $(X_i - \bar{X})/S$ ***en lugar del valor en la tabla*** τ***?***

Si bien para cada caso particular es posible mejorar marginalmente la predicción utilizando un valor de corte fijo específico, no encontramos uno para el cual la predicción mejore en todos los casos. Esto se debe en parte a que las rutas no sólo son de largos distintos sino que también sufren de anomalías distintas. Además, los servicios de geolocalización pueden estar desactualizados, todo esto resulta en que cada caso sea distinto y difícil de generalizar.

4. Conclusión

Se realizaron experimentos de traceroutes a 4 distintos destinos a lo largo del mundo con el fin de analizar sus características, comportamiento y anomalías. Se pudo identificar la ubicación de los routers que componen estas rutas para contrastar con los RTT obtenidos. Las considerables anomalías resultaron un importante impedimento a la hora de extraer información de los resultados si no se complementase éstos con datos de geolocalización.

Se analizó la eficiencia del método Thompson tau modificado para determinar saltos interoceánicos de las rutas. Este tuvo un return decente pero con el costo de numerosos falsos positivos.

¿Qué porcentaje de saltos no responden los Time exceeded? ¿Cuál es el largo de la ruta en términos de los saltos que sí responden?

- Australia: Responden a time exceeded 79 % - Longitud de la ruta: 24 saltos.
- India: Responden a time exceeded 82 % - Longitud de la ruta: 17 saltos.
- Qatar: Responden a time exceeded 83 % - Longitud de la ruta: 12 saltos.
- Sudafrica: Responden a time exceeded 83 % - Longitud de la ruta: 18 saltos.

¿La ruta tiene enlaces intercontinentales? ¿Cuántos?

Elegimos nuestras rutas de manera que se encuentren a por lo menos un salto interoceánico de distancia (Todos nuestros destinos están fuera de América), es por eso que la respuesta es sí para todas las rutas necesariamente. Algunas rutas, sin embargo, resultaron atravesar océanos más de una vez, como es el caso de Sudáfrica.

- Australia: 1 salto intercontinental
- India: 2 saltos intercontinentales
- Qatar: 1 salto intercontinental
- Sudafrica: 3 saltos intercontinentales

¿Se observaron comportamientos anómalos del tipo descripto en la bibliografía sugerida?

Sí, obtuvimos RTTs no crecientes para TTLs crecientes como se puede observar en el cuadro 3. El paquete con TTL = 6 tomó $42ms$ en responder mientras que el TTL=7 hizo un round trip en bastante menos, $20ms$. Estos no fueron los resultados de una iteración, si no el promedio de 30 de ellas, en todos los casos donde 181.96.113.234 respondió, lo hizo tardando más que lo que tardó 195.22.220.56 en hacerlo, un paso más lejos en la ruta. También vale aclarar que mientras que 195.22.220.56 respondió 30 de los 30 paquetes ICMP, 181.96.113.234 respondió tan solo 6 de ellos.

Esto nos indica que el TTL = 6 probablemente tenga un balanceador de carga por delante, donde 181.96.113.234 es uno de los posibles routers, uno particularmente lento. Y el camino al router 195.22.220.56 pasa por otro router que no responde a paquetes ICMP.

Esta misma situacion se puede observar en el cuadro **??** donde la misma IP se comporta de la misma manera. En el caso de India la diferencia es tan grande que de incluirla como dato válido produciría cambios drásticos en nuestro análisis tanto de outliers y saltos interoceánicos como en nuestro cáculo de diferencia de RTT entre routers.

Otra anomalía que a priori se ve similar en naturaleza es lo que sucede en los últimos paquetes ICMP del cuadro 3. Para TTL = 15 observamos un RTT de 455ms mientras que para TTL = 18 (más adelante en la ruta) el RTT es de 410ms. Esto, a pesar de parecer similar a las anomalías observadas previamente, nos sugiere algo diferente ya que tanto TTL = 15 como TTL = 17 fueron respondidos 30 de 30 intentos. Esto nos sugiere que probablemente el camino de ida no sea igual al camino de vuelta, cosa que tiene sentido teniendo en cuenta que se trata de routers muy lejanos en cantidad de saltos y nos recuerda nuevamente lo limitado que es este método para entender la topología de la red.

¿Se observaron otros comportamientos anómalos? Proponga hipótesis que permitan explicarlos.

No se observaron otros comportamientos anómalos.

¿Se aprecia alguna diferencia en la capacidad de detectar enlaces intercontinentales según el largo de la ruta?

No, en todas los tracerutes se obtuvo un desempeño similar a la hora de detectar saltos intercontinentales, no detectándose tan solo un salto (India y Sudáfrica) o detectándose todos (Australia y Qatar).

¿Los traceroutes recorren la mínima distancia posible? ¿Hay una correlación entre distancia geográfica y la cantidad de saltos?

No, como se puede ver claramente en la 3 ningún traceroute realiza un recorrido eficiente, ni siquiera la ruta a la India, que es la que menos saltos realiza, es eficiente debido a que los paquetes primero viajan a Córdoba para después volver a Buenos Aires para finalmente salir a Roma. Nuestra hipótesis es que esto se debe a que el protocolo IP es best effort, por lo que los paquetes hacen su "mejor intento.[en] llegar a destino, por lo que a menudo toman rutas equivocadas o subóptimas. El análisis de este fenómeno escapa el alcance de este informe.

5. Anexo

Referencias

[1] Código utilizado para realizar los experimentos:
`https://github.com/vicentevieytes/TDLC-TP2-Traceroute`

[2] Geolocalizadores de IP utilizados:
`https://ipapi.co/`
`https://ipinfo.io/tools/map`
`https://www.iplocation.net/ip-lookup`
`https://ipgeolocation.io/`

[3] Martin Erich Jobst. (2012). *Traceroute Anomalies* Universidad Técnica de Múnich.
`https://www.net.in.tum.de/fileadmin/TUM/NET/NET-2012-08-1/NET-2012-08-1_02.pdf`

[4] John M. Cimbala. (2011). *Outliers* Penn State University.
`https://www.me.psu.edu/cimbala/me345/Lectures/Outliers.pdf`