



DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico: Rutas en Internet

20 de Octubre de 2019

Teoría de las Comunicaciones

Integrante	LU	Correo electrónico
Sebastián Fernández Ledesma	392/06	sfernandezledesma@gmail.com
Sebastián Garbi	179/05	garbyseba@gmail.com
Durán, Alejandro	286/05	alejandrofduran@gmail.com



Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Ciudad Universitaria - Pabellón I

Intendente Güiraldes 2160 - C1428EGA

Ciudad Autónoma de Buenos Aires - Argentina

Tel/Fax: (54 11) 4576-3359

<http://exactas.uba.ar>

Índice

1. Introducción	3
2. Métodos	3
2.1. Análisis general de una ruta	3
2.2. Detección de enlaces intercontinentales	3
3. Resultados	5
3.1. Italia (Sant’Anna School of Advanced Studies)	5
3.1.1. Información de la ruta promedio	5
3.1.2. Mapeo de la ruta en el planisferio	5
3.1.3. Enlaces intercontinentales	6
3.1.4. Gráficos de RTT entre saltos	7
3.2. Japón (Aichi Bunkyo University)	8
3.2.1. Información de la ruta promedio	8
3.2.2. Mapeo de la ruta en el planisferio	8
3.2.3. Enlaces intercontinentales	10
3.2.4. Gráficos de RTT entre saltos	11
3.3. Sudáfrica (University of South Africa)	12
3.3.1. Información de la ruta promedio	12
3.3.2. Mapeo de la ruta en el planisferio	12
3.3.3. Enlaces intercontinentales	13
3.3.4. Gráficos de RTT entre saltos	14
4. Conclusiones	15

1. Introducción

En este trabajo analizaremos rutas en internet con la intención de hallar enlaces intercontinentales. Para ello implementaremos una herramienta similar a **traceroute** y experimentaremos con ella analizando las rutas a distintas universidades de otros continentes, y estimaremos cuáles enlaces son intercontinentales basándonos en el algoritmo propuesto por Cimbala¹.

2. Métodos

2.1. Análisis general de una ruta

El análisis de ruta consiste en utilizar los mensajes *TimeExceeded* de respuesta de los routers intermedios para encontrar los saltos de una ruta hacia el destino que buscamos.

El mecanismo consiste en enviar paquetes **ICMP EchoRequest** con **TTL** desde 1 a un número máximo hasta encontrar el host de destino. De este modo con **TTL** = 1 el primer router que lo agarre decrementará el **TTL** y al ser 0 lo descartará enviando el *TimeExceeded* correspondiente con la IP del router que lo decrementó. De esta manera, armaremos la ruta, incrementando el TTL hasta que la ip que responda sea el destino original o se alcance el máximo TTL.

Cabe aclarar que ésta metodología es una heurística ya que los envíos con distinto **TTL** son independientes unos de otros y podrían llegar a tomar caminos distintos y generar links entre routers inexistentes².

Algunos problemas con los que nos encontramos utilizando este método, es que no todos los routers responden las *EchoRequest* por cuestiones de seguridad, dejando varios saltos desconocidos en la ruta. Modificamos el mecanismo para que cuando no obtenga una respuesta a la petición **ICMP** envíe una nueva petición utilizando *UDP* con puerto 53, y de esta manera pudimos mejorar bastante la cantidad de nodos que responden.

2.2. Detección de enlaces intercontinentales

Para la estimación de los enlaces intercontinentales, nuestra herramienta se vale de repetir una cantidad n de traceroutes (que a su vez se valen del programa *scapy* para hacer los pings y obtener las respuestas), y de promediar las restas de RTT entre saltos de TTL contiguo (esto es equivalente a promediar los RTT para cada TTL y luego hacer las restas correspondientes). Las rutas con las que experimentamos resultaron ser muy estables, las pocas veces que se toma una ruta alternativa se descartan por ser despreciables (pidiendo una cota de mediciones mínimas para ser consideradas). Luego, usando el método de Cimbala¹ con una modificación (que detallaremos luego), se buscan outliers de estos valores, que serán los saltos intercontinentales estimados.

Inicialmente, experimentamos con todos los tiempos entre saltos y usando el método original de Cimbala, pero rápidamente hallamos que había serios problemas tanto en las mediciones como en el método de búsqueda de outliers:

- Por un lado, encontramos anomalías² en los RTT de cada salto, esto es, el RTT de un salto con TTL i era mayor que el de TTL $i + 1$, lo cual incluso al promediar terminaba dando un valor negativo. Esta anomalía puede deberse a que los caminos de regreso entre un router y el siguiente pueden ser diferentes, y por lo tanto ser el segundo más rápido que el primero, dando resultados inconsistentes. Como este tipo de valores no son válidos, decidimos no considerarlos.
- El segundo problema es la enorme cantidad de falsos positivos, algunos absurdos, como tiempos entre routers de un milisegundo. La causa de ésto es múltiple:
 - El método de Cimbala busca outliers *en general*, y los candidatos serán siempre los máximos o los mínimos de la muestra. Pero nosotros sabemos que un mínimo no tiene sentido que sea un enlace intercontinental.

¹<http://www.mne.psu.edu/cimbala/me345/Lectures/Outliers.pdf>

²http://www.net.in.tum.de/fileadmin/TUM/NET/NET-2012-08-1/NET-2012-08-1_02.pdf

- La varianza es muy alta, y el método de Cimballa va sacando de a un outlier y recalculando la media y el desvío estándar para la muestra sin el outlier, pero incluso sacando los enlaces intercontinentales, la varianza es tal que se siguen hallando outliers que claramente no son lo que buscamos (con valores por ejemplo de 29 ms).

Para resolver ambos problemas usamos la siguiente heurística: los candidatos a outliers deben a su vez ser mayores que la **media total** de *todos* los tiempos entre los nodos. Experimentalmente vimos que la heurística funciona muy bien, ya que los enlaces intercontinentales tienen tiempos muy mayores al resto, y siempre están por encima de la media total.

3. Resultados

3.1. Italia (Sant'Anna School of Advanced Studies)

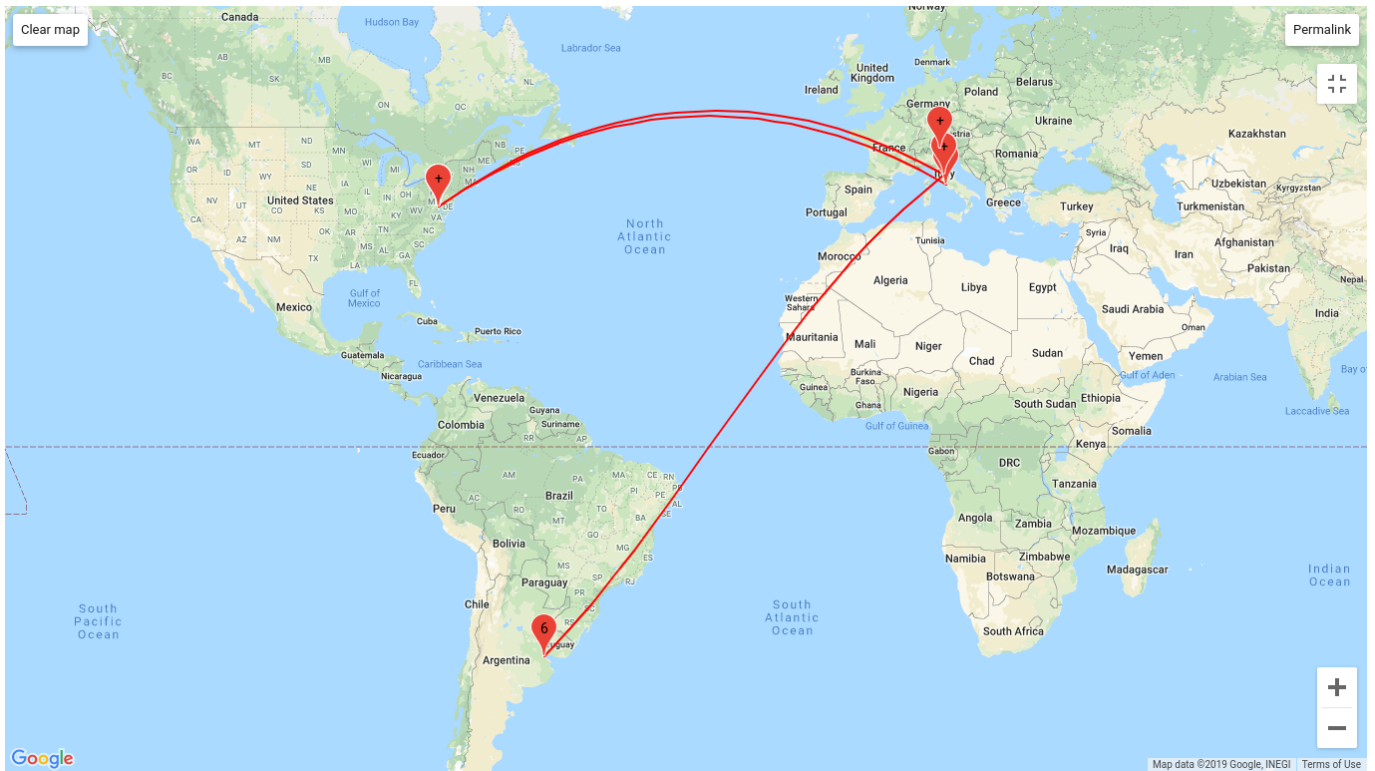
3.1.1. Información de la ruta promedio

Notar que según la geolocalización, se va hasta Italia, luego a Estados Unidos, y se vuelve a Europa. Hablaremos de esta anomalía en el análisis de los gráficos de la ruta en el planisferio.

TTL	RUTA	\overline{RTT}	$\overline{RTT_i} - \overline{RTT_{i-1}}$	PAÍS
1	192.168.0.1	80,00	80,00	IP privada (Red Privada)
2				
3				
4				
5				
6	200.89.165.250	87,60	87,60	Argentina
7	185.70.203.32	76,70	-10,90	Italia
8	89.221.41.181	211,40	134,70	Italia
9	154.54.9.17	187,00	-24,40	Estado Unidos
10	154.54.47.17	180,50	-6,50	Estado Unidos
11	154.54.24.145	208,40	27,90	Estado Unidos
12	154.54.7.157	205,80	-2,60	Estado Unidos
13	154.54.40.105	217,40	11,60	Estado Unidos
14	66.28.4.198	295,90	78,50	Francia
15	130.117.49.154	312,70	16,80	Francia
16	130.117.1.42	307,50	-5,20	Italia
17	154.54.57.65	321,60	14,10	Italia
18	149.6.22.74	306,80	-14,80	Italia
19	90.147.80.206	319,20	12,40	Italia
20	193.206.136.46	311,10	-8,10	Italia
21	193.205.80.112	321,50	10,40	Italia

3.1.2. Mapeo de la ruta en el planisferio

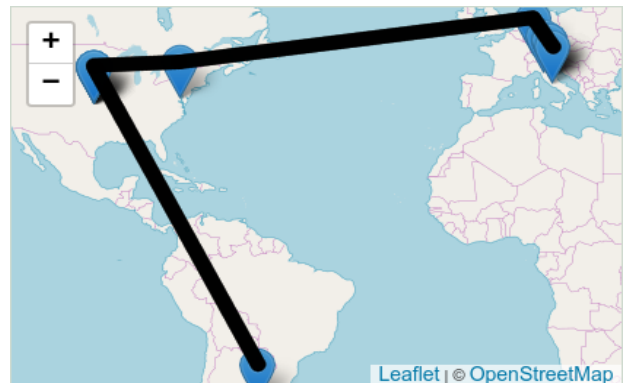
Aquí encontramos dificultades con las páginas que geolocalizan direcciones IP. Para empezar, no siempre se detectan las mismas localizaciones para la misma IP, pero aún más grave es que aparentemente algunas IP son localizadas de manera absurda. Por ejemplo, la ruta hacia Italia en obtenida usando <https://stefansundin.github.io/traceroute-mapper/> es la siguiente:



Como se puede observar, para llegar Italia, se llega efectivamente en uno de los saltos, para luego cruzar el Atlántico, pasar por algunos routers en Estados Unidos, y luego cruzar de nuevo hasta Europa. Esto claramente no tiene sentido, ya que se tarda en promedio 134 ms para ir de Argentina a Italia, y aparentemente en un instante se pasa por routers estadounidenses. Nuestra teoría es que las IP están mal asignadas a Italia, quizás porque la empresa dueña de los routers es Telecom Italia. En lugar de la ruta sugerida por la geolocalización, creemos que la más coherente es la que podemos ver a continuación donde se la compara con la original:



(a) Ruta según geolocalización



(b) Probablemente la verdadera ruta

Figura 1: Comparación de rutas posibles

3.1.3. Enlaces intercontinentales

Output de la herramienta para 60 mediciones:

TTL1	TTL2	IP1	IP2	Tiempo entre nodos
7	8	185.70.203.32	89.221.41.181	149 ms
10	11	154.54.47.17	154.54.24.145	30 ms
12	13	154.54.7.157	154.54.40.105	3 ms
13	14	154.54.40.105	66.28.4.198	74 ms
14	15	66.28.4.198	130.117.49.154	21 ms
16	17	130.117.1.42	154.54.57.65	10 ms
18	19	149.6.22.74	90.147.80.206	9 ms

19	20	90.147.80.206	193.206.136.46	0 ms
TTL1	TTL2	IP1	IP2	Valor Z
7	8	185.70.203.32	89.221.41.181	2.3445537987722576
10	11	154.54.47.17	154.54.24.145	-0.1465346124232661
12	13	154.54.7.157	154.54.40.105	-0.7117395460558639
13	14	154.54.40.105	66.28.4.198	0.7745400942372637
14	15	66.28.4.198	130.117.49.154	-0.3349362569674654
16	17	130.117.1.42	154.54.57.65	-0.5652049336325978
18	19	149.6.22.74	90.147.80.206	-0.5861384496930644
19	20	90.147.80.206	193.206.136.46	-0.7745400942372637

OUTLIERS USANDO CIMBALA:

TTL1	TTL2	IP1	IP2	Tiempo entre nodos
7	8	185.70.203.32	89.221.41.181	149 ms
13	14	154.54.40.105	66.28.4.198	74 ms

OUTLIERS PARA RESULTADOS NORMALIZADOS:

TTL1	TTL2	IP1	IP2	Valor Z
7	8	185.70.203.32	89.221.41.181	2.3445537987722576
13	14	154.54.40.105	66.28.4.198	0.7745400942372637

Como mencionamos antes, sólo consideramos los tiempos entre routers mayores o iguales a cero milisegundos, para no incluir anomalías en los cálculos. También vemos los valores normalizados, que no añaden información y dan los mismos outliers. En este caso, encontramos dos enlaces intercontinentales, uno entre Sudamérica y Estados Unidos, y el otro entre Estados Unidos y Europa.

3.1.4. Gráficos de RTT entre saltos

Aquí podemos observar cómo los valores de los dos enlaces intercontinentales están por encima de la media total de los tiempos entre routers, y cómo al pedir que los candidatos a outliers estén por encima, evitamos falsos positivos como el tiempo de 30 ms.

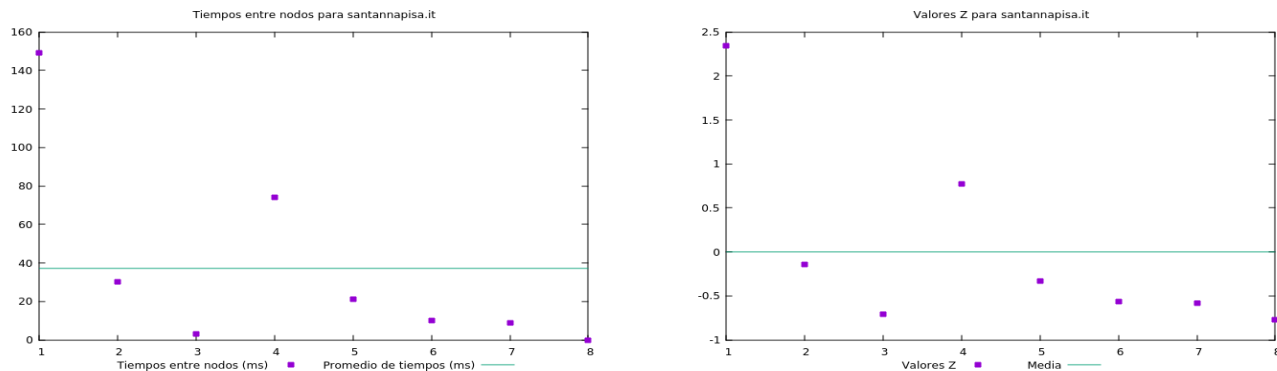


Figura 2: Tiempos promediados entre routers y valores normalizados

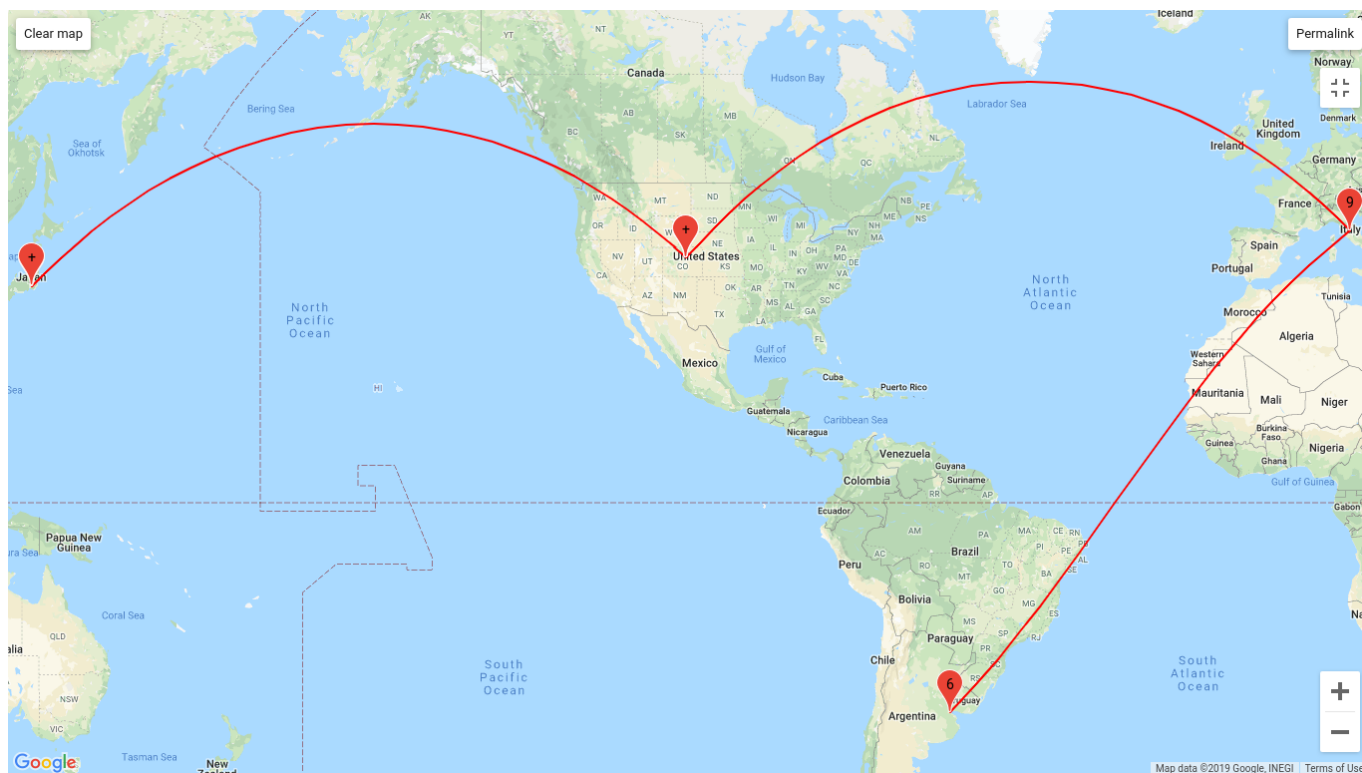
3.2. Japón (Aichi Bunkyo University)

3.2.1. Información de la ruta promedio

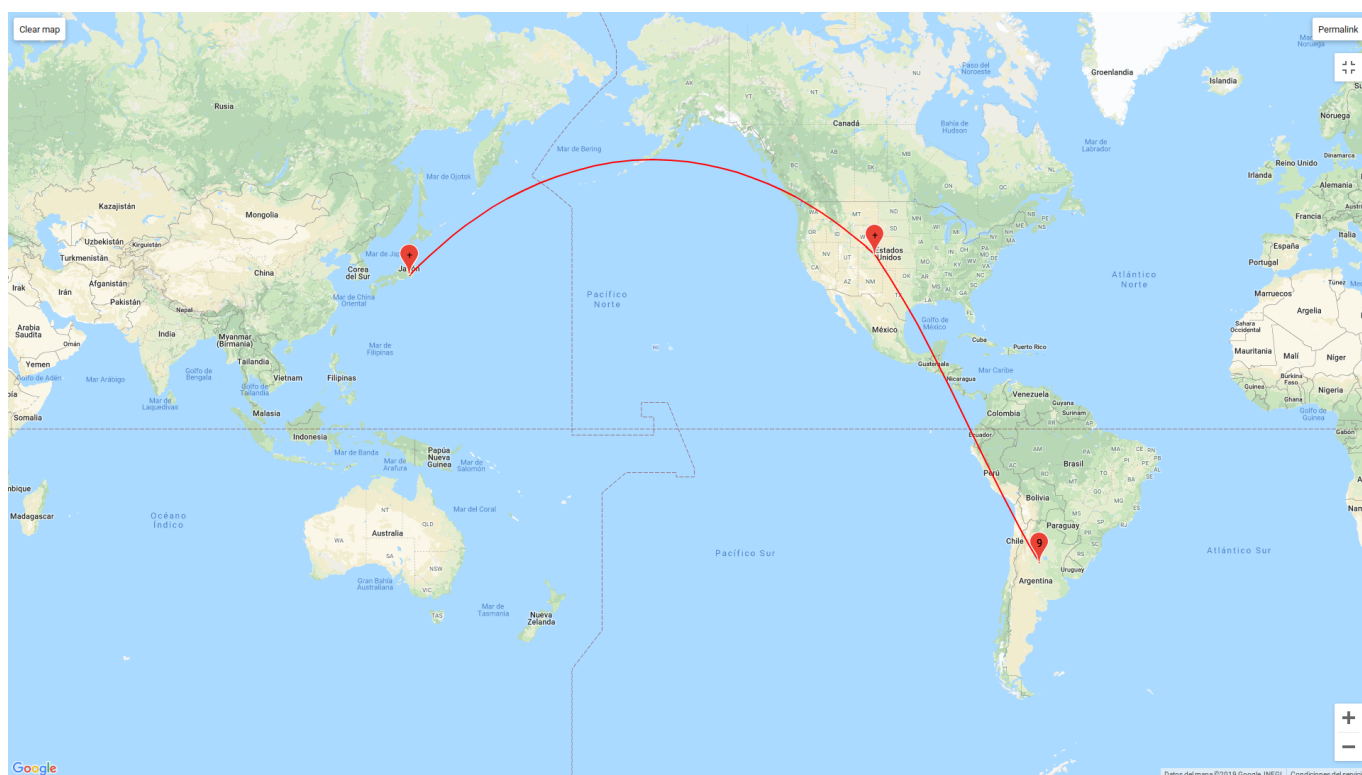
TTL	IP	\overline{RTT}	$\overline{RTT}_i - \overline{RTT}_{i-1}$	PAÍS
1	192.168.0.1	57,27	57,27	IP privada (Red Privada)
2	10.33.64.1	59,47	2,20	IP privada (Proveedor de Servicio)
3	100.72.4.149	61,80	2,33	No Disponible
4	100.72.3.194	58,40	-3,40	No Disponible
5	100.72.3.193	65,80	7,40	No Disponible
6	100.72.3.185	55,00	-10,80	No Disponible
7	10.242.5.18	59,27	4,27	IP privada (Proveedor de Servicio)
8	200.32.127.98	60,27	1,00	Argentina, Buenos Aires
9	200.32.127.97	91,13	30,86	Argentina, Buenos Aires
10	67.17.(99.233 94.249)	183,67	92,54	Estados Unidos
11	*			
12	*			
13	4.68.127.54	193,27	193,27	Estados Unidos
14	129.250.(2.202 3.40)	337,93	144,66	Estados Unidos
15	129.250.4.(20 88)	197,36	-140,57	Estados Unidos
16	129.250.2.219	211,27	13,91	Estados Unidos
17	129.250.7.69	234,93	23,66	Japón
18	129.250.2.177	332,27	97,34	Japón
19	129.250.(7.33 2.128)	332,07	-0,20	Japón
20	129.250.3.(106 232 210 88)	344,73	12,67	Japón
21	61.200.(82.178 91.186)	322,71	-22,02	Japón
22	*			
23	*			
24	*			
25	210.188.213.76	343,27	343,27	Japón
26	183.90.238.55	340,13	-3,14	Japón

3.2.2. Mapeo de la ruta en el planisferio

Nuevamente, la geolocalización de la ruta no tiene demasiado sentido ya que salta hacia Europa para despues ir a Estados Unidos.



Analizando un poco mejor el caso, notamos que el problema se origina para los ttl 10 a 13 los cuales según el sitio que genera el mapa³ ubica esos ips en Suiza. Evaluando puntualmente esas ips en otro sitio de geolocalización que obtiene resultados de 3 fuentes distintas⁴ obtenemos que en 2 de los 3 resultados ubica los ips en Estados unidos y sólo uno en Suiza. Removiendo los ips problematicos obtenemos un mapa un poco más coherente.



³<https://stefansundin.github.io/traceroute-mapper/>

⁴<https://www.iplocation.net/>

3.2.3. Enlaces intercontinentales

Output de la herramienta para 60 mediciones:

TTL1	TTL2	IP1	IP2	Tiempo entre nodos
1	2	192.168.0.1	10.33.64.1	2 ms
2	3	10.33.64.1	100.72.4.149	2 ms
4	5	100.72.3.194	100.72.3.193	7 ms
6	7	100.72.3.185	10.242.5.18	4 ms
7	8	10.242.5.18	200.32.127.98	1 ms
8	9	200.32.127.98	200.32.127.97	29 ms
13	14	4.68.127.54	129.250.2.202	148 ms
15	16	129.250.4.20	129.250.2.219	13 ms
16	17	129.250.2.219	129.250.7.69	23 ms
17	18	129.250.7.69	129.250.2.177	98 ms
18	19	129.250.2.177	129.250.7.33	4 ms
13	14	4.68.127.54	129.250.3.40	143 ms
TTL1	TTL2	IP1	IP2	Valor Z
1	2	192.168.0.1	10.33.64.1	-0.695728049086611
2	3	10.33.64.1	100.72.4.149	-0.695728049086611
4	5	100.72.3.194	100.72.3.193	-0.6029643092083962
6	7	100.72.3.185	10.242.5.18	-0.6586225531353251
7	8	10.242.5.18	200.32.127.98	-0.714280797062254
8	9	200.32.127.98	200.32.127.97	-0.19480385374425108
13	14	4.68.127.54	129.250.2.202	2.0129731553572614
15	16	129.250.4.20	129.250.2.219	-0.49164782135453844
16	17	129.250.2.219	129.250.7.69	-0.30612034159810886
17	18	129.250.7.69	129.250.2.177	1.0853357565751132
18	19	129.250.2.177	129.250.7.33	-0.6586225531353251
13	14	4.68.127.54	129.250.3.40	1.9202094154790463

OUTLIERS USANDO CIMBALA:

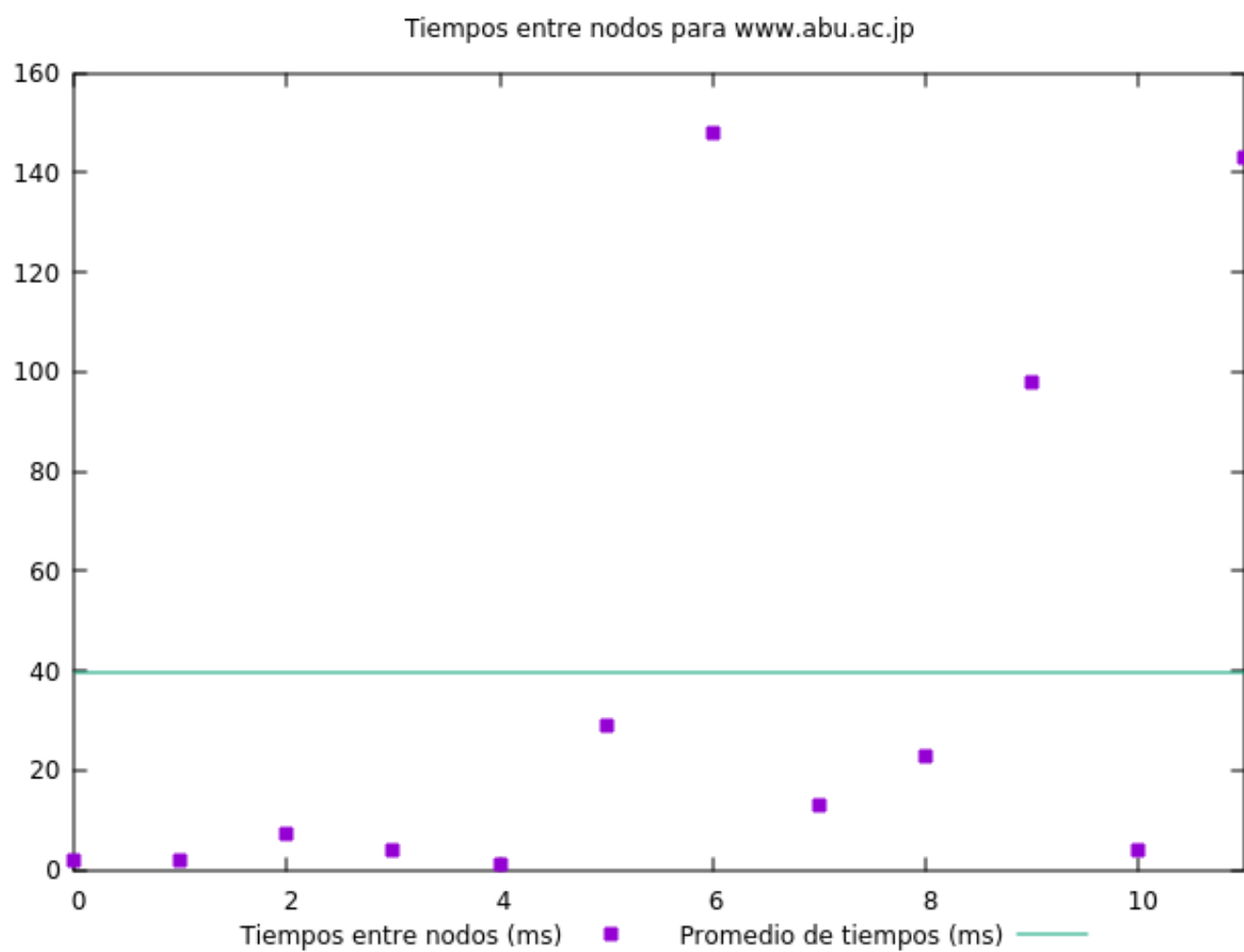
TTL1	TTL2	IP1	IP2	Tiempo entre nodos
13	14	4.68.127.54	129.250.2.202	148 ms
13	14	4.68.127.54	129.250.3.40	143 ms
17	18	129.250.7.69	129.250.2.177	98 ms

OUTLIERS PARA RESULTADOS NORMALIZADOS:

TTL1	TTL2	IP1	IP2	Valor Z
13	14	4.68.127.54	129.250.2.202	2.0129731553572614
13	14	4.68.127.54	129.250.3.40	1.9202094154790463
17	18	129.250.7.69	129.250.2.177	1.0853357565751132

En este caso, encontramos dos enlaces intercontinentales, uno entre Sudamérica y Estados Unidos, y el otro entre Estados Unidos y Japón. Cabe destacar que para el caso de ttl 13 a 14 se encontraron 2 posibles enlaces, por lo que suponemos que podría llegar a haber al menos dos enlaces intercontinentales paralelos en ese salto.

3.2.4. Gráficos de RTT entre saltos



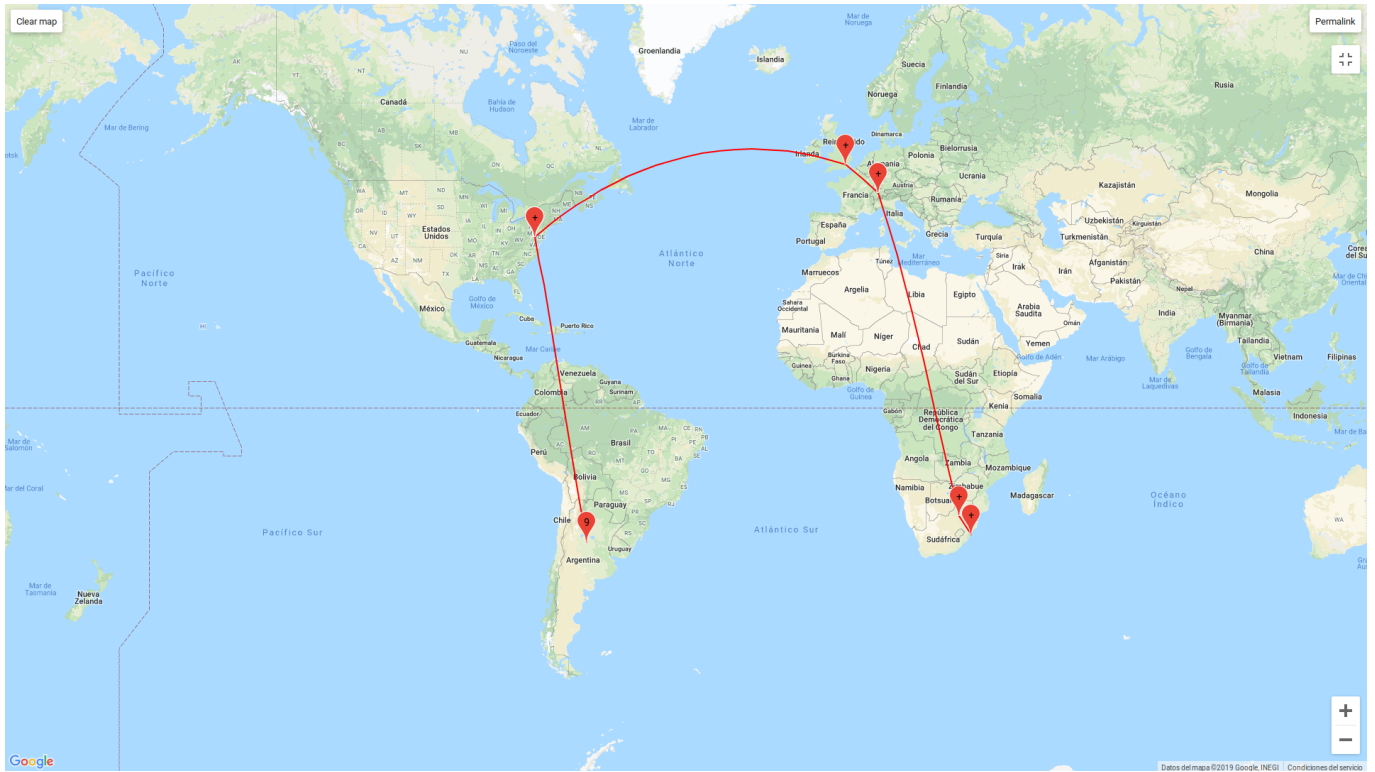
3.3. Sudáfrica (University of South Africa)

3.3.1. Información de la ruta promedio

1	192.168.0.1	49,40	49,40	IP privada (Red Privada)
2	10.33.64.1	51,20	1,80	IP privada (Proveedor de Servicio)
3	100.72.4.149	52,60	1,40	No Disponible
4	100.72.3.194	64,80	12,20	No Disponible
5	100.72.3.193	63,00	-1,80	No Disponible
6	100.72.3.185	52,20	-10,80	No Disponible
7	10.242.5.18	63,80	11,60	IP privada (Proveedor de Servicio)
8	200.32.127.98	62,20	-1,60	Argentina, Buenos Aires
9	200.32.127.97	54,00	-8,20	Argentina, Buenos Aires
10	67.17.99.233	184,60	130,60	Estados Unidos
11				
12				
13	154.54.10.57	186,40	186,40	Estados Unidos
14	154.54.47.(29 17)	175,80	-10,60	Estados Unidos
15	154.54.(28.49 24.145)	192,00	16,20	Estados Unidos
16	154.54.(24.221 7.157)	209,60	17,60	Estados Unidos
17	154.54.40.(109 105)	205,60	-4,00	Estados Unidos
18	154.54.(42.86 30.186)	265,80	60,20	Estados Unidos
19	154.54.(58.186 57.154)	269,00	3,20	Estados Unidos
20	154.54.61.61	263,00	-6,00	Estados Unidos
21	149.14.146.194	269,00	6,00	Reino Unido
22	155.232.1.40	406,80	137,80	Sudafrica
23	155.232.64.76	444,20	37,40	Sudafrica
24	155.232.1.(34 66)	442,00	-2,20	Sudafrica
25	155.232.15.(105 212)	446,40	4,40	Sudafrica
26	155.232.29.2	438,80	-7,60	Sudafrica

3.3.2. Mapeo de la ruta en el planisferio

Esta vez, la geolocalización dió resultados coherentes, se pueden observar los tres enlaces intercontinentales que hallaremos con la herramienta.



3.3.3. Enlaces intercontinentales

Output de la herramienta para 60 mediciones:

TTL1	TTL2	IP1	IP2	Tiempo entre nodos
7	8	185.70.203.32	89.221.41.171	143 ms
9	10	154.54.9.17	154.54.47.29	2 ms
10	11	154.54.47.29	154.54.28.49	15 ms
11	12	154.54.28.49	154.54.24.221	7 ms
12	13	154.54.24.221	154.54.40.109	5 ms
13	14	154.54.40.109	154.54.42.86	82 ms
16	17	154.54.61.69	149.14.146.194	0 ms
17	18	149.14.146.194	155.232.1.40	150 ms
18	19	155.232.1.40	155.232.64.76	19 ms
19	20	155.232.64.76	155.232.1.16	11 ms
20	21	155.232.1.16	155.232.15.214	0 ms
TTL1	TTL2	IP1	IP2	Valor Z
7	8	185.70.203.32	89.221.41.171	1.880957264023117
9	10	154.54.9.17	154.54.47.29	-0.6803813808406709
10	11	154.54.47.29	154.54.28.49	-0.44422959088869046
11	12	154.54.28.49	154.54.24.221	-0.5895537693206784
12	13	154.54.24.221	154.54.40.109	-0.6258848139286753
13	14	154.54.40.109	154.54.42.86	0.7728604034792087
16	17	154.54.61.69	149.14.146.194	-0.7167124254486679
17	18	149.14.146.194	155.232.1.40	2.0081159201511065
18	19	155.232.1.40	155.232.64.76	-0.3715675016726965
19	20	155.232.64.76	155.232.1.16	-0.5168916801046844
20	21	155.232.1.16	155.232.15.214	-0.7167124254486679

OUTLIERS USANDO CIMBALA:

TTL1	TTL2	IP1	IP2	Tiempo entre nodos
17	18	149.14.146.194	155.232.1.40	150 ms
7	8	185.70.203.32	89.221.41.171	143 ms
13	14	154.54.40.109	154.54.42.86	82 ms

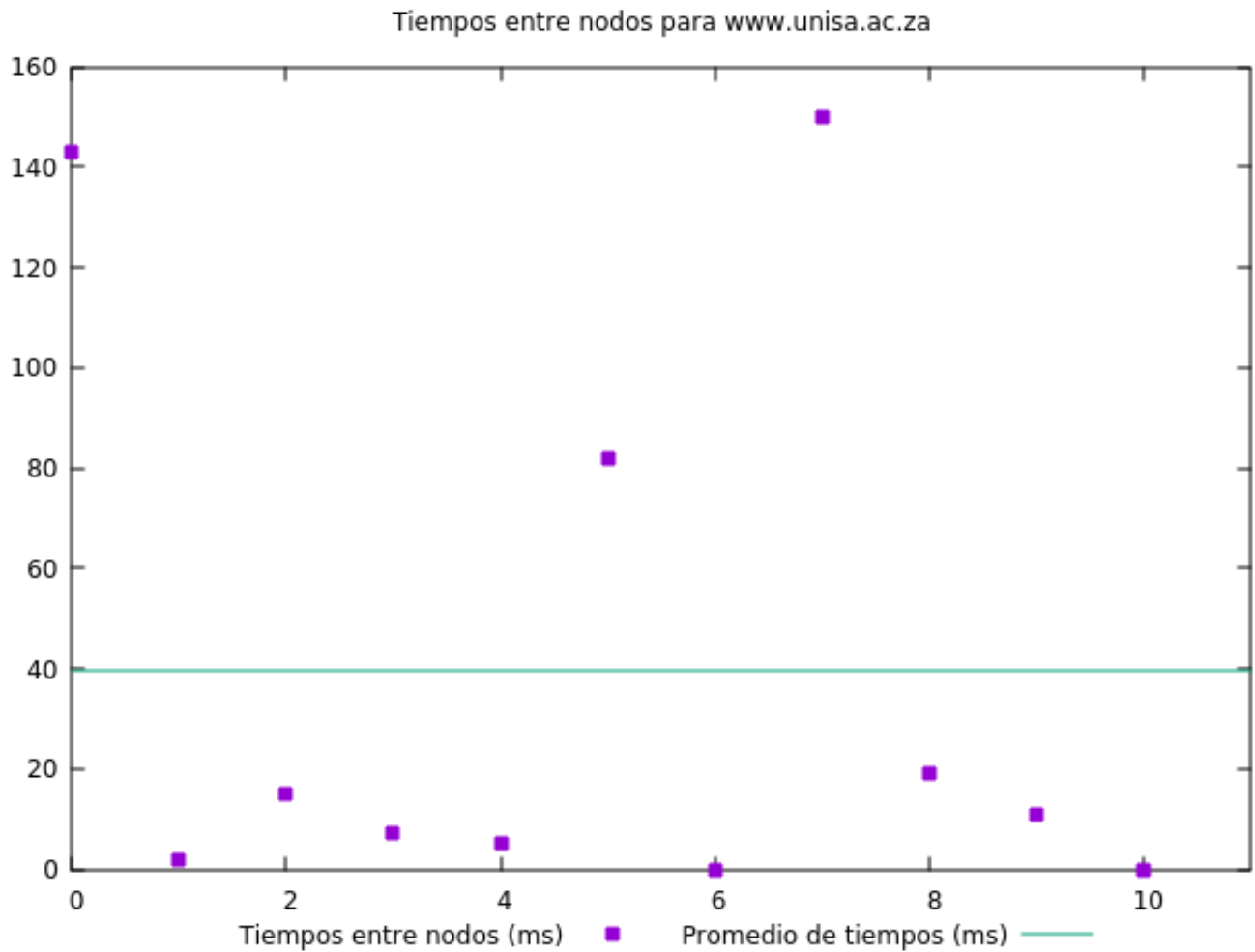
OUTLIERS PARA RESULTADOS NORMALIZADOS:

TTL1	TTL2	IP1	IP2	Valor Z
17	18	149.14.146.194	155.232.1.40	2.0081159201511065
7	8	185.70.203.32	89.221.41.171	1.880957264023117
13	14	154.54.40.109	154.54.42.86	0.7728604034792087

La herramienta logró encontrar correctamente los tres enlaces intercontinentales.

3.3.4. Gráficos de RTT entre saltos

De nuevo, vemos como los valores de los enlaces intercontinentales están por encima de la media total.



4. Conclusiones

En este trabajo analizamos rutas y estimamos enlaces intercontinentales. Pudimos observar varias anomalías, la más importante de ellas siendo los tiempos negativos entre routers, que en principio pensamos que eran por la varianza de tiempos, pero se mantuvieron incluso aumentando el número de mediciones o tomando los mínimos de los RTT de los saltos en lugar de los promedios. Eliminando estos valores y con la modificación al método de Cimbala obtuvimos buenas estimaciones de cuáles son los enlaces intercontinentales.

A pesar de ello, un punto débil pueden ser las rutas que no posean estos enlaces, ya que al haber tanta varianza en las muestras, la herramienta desarrollada puede dar falsos positivos, por lo cual una idea es usar una cota fija para los candidatos a outlier en lugar de nuestra heurística.