Detección de Enlaces Intercontinentales

Manuel Costa Mathias Gatti

Resumen —

Index Terms —**traceroute, enlaces intercontinentales, ICMP, anomalías, RTT, TTL**

I. Introducción

— COMPLETAR —

Iniciamos este trabajo con el objetivo de aprender sobre el protocolo ARP y más en general las redes de Internet y sus algoritmos de intercambio de paquetes.

A lo largo de este informe describiremos ciertos análisis que tendrán como objetivo el modelado de emisores de información y la posterior detección de símbolos destacados. A partir de esta metodología creemos que podremos identificar dispositivos claves en la red como por ejemplo el gateway. Nuestra hipótesis es que este tendrá destacará en el intercambio de paquetes who-has siendo de los que mas reciba y envíe.

— COMPLETAR —

II. Métodos

Herramientas

Para la implementación de traceroute utilizamos el código provisto por la catedra al cual le realizamos ciertas modificaciones para poder detectar anomalías y guardar los datos obtenidos de forma más cómoda.

Detección de Anomalías

Como se describe en Cimbala[1] aplicamos un cálculo el cual compara los datos normalizados con z-score con una valor τ derivado del t-student con un α de 0.05 lo cual modela un intervalo de confianza el cual descarta los 0.025 percentiles.

Capturas

Se corrió traceroute sobre 3 universidades distintas con un ttl de 30 y 40 queries. Con el objetivo de lograr contrastar elegimos universidades muy lejanas y muy cercanas. A continuación describimos brevemente a cada una.

- Universidad de São Paulo (www5.usp.br) esta será la universidad mas cercana, ubicada en el mismo continente, por lo cual esperamos que no haya ningún enlace intercontinental.
- Universidad de Sidney (www.sydney.edu.au) escogimos esta universidad ya que nos surgió la duda de si existirá algún enlace intercontinental directo entre oceanía y america o tendrá que pasar europa resultando en varios enlaces.
- Universidad de Moscú (www.msu.ru) al estar ubicada en un punto tan alejado de nosotros estabamos seguros de que iba a haber algún salto intercontinental y quizás más.

M. Costa, e-mail: manucos94@gmail.com
M. Gatti, e-mail: mathigatti@gmail.com

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

III. Resultados y análisis

Los resultados de cada ruta se presentan en dos partes: primero se realiza una descripción geográfica de la ruta trazada, basada en las ips seguidas, prestando particular atención a los saltos intercontinentales; en la segunda parte se analizan los RTTs entre saltos, de los cuales se busca poder inferir en forma automática los enlaces intercontinentales detectados en la primer parte.

Universidad de São Paulo

Recorrido en el Planisferio

COMPLETAR: En esta parte creo que deberiamos responder las preguntas ¿Que porcentaje de saltos no responden los Time exceeded? ¿Cual es el largo de la ruta en terminos de los saltos que si responden? (todo esto se puede sacar viendo los csv) ¿La ruta tiene enlaces intercontinentales? ¿Cuantos?

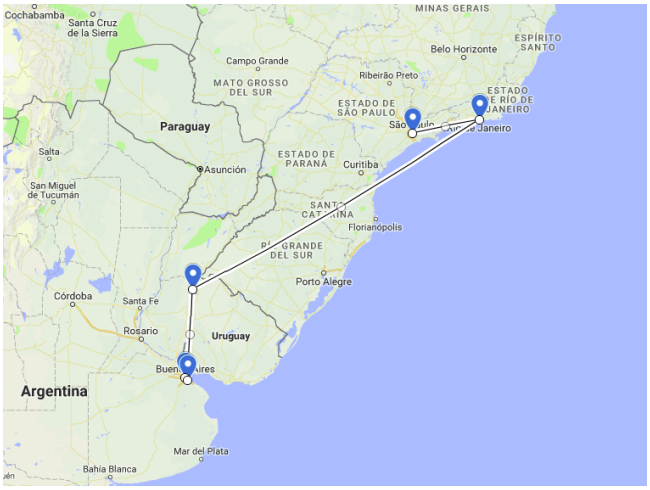

Figura 1: Recorrido realizado por los paquetes durante la ejecución de traceroute al intentar alcanzar el sitio `www5.usp.br`

RTT entre saltos

Analicemos cómo funciona nuestro modelo para inferir saltos intercontinentales, en un caso donde sabemos que no hay ninguno.

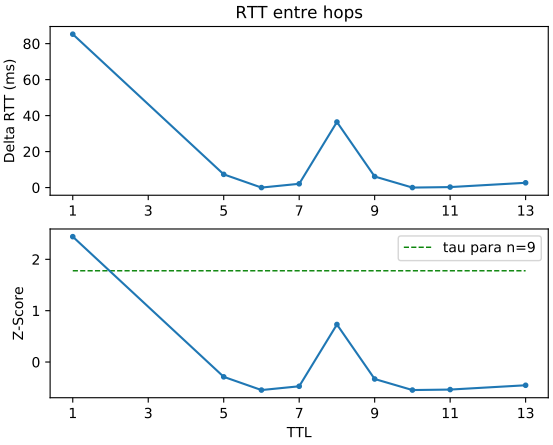

Figura 2: RTT entre saltos (antes y después de normalizar respectivamente) para el sitio `www5.usp.br`. n es la cantidad de TTLs para los que se obtuvo un *time exceed* (que son los puntos que se grafican).

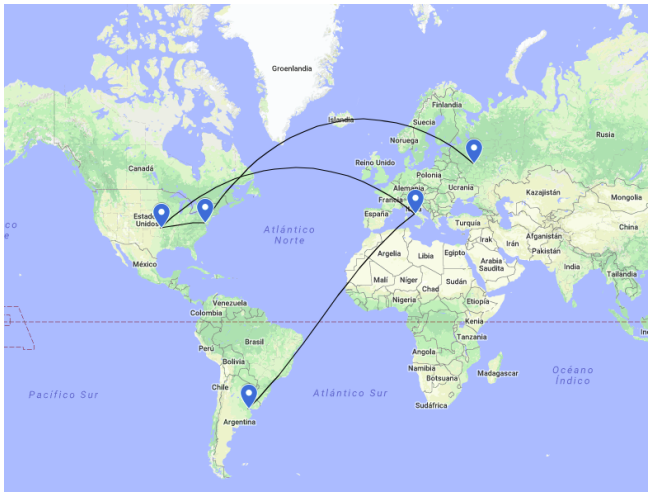

Figura 3: Recorrido realizado por los paquetes durante la ejecución de traceroute al intentar alcanzar el sitio `www.msu.com`

Viendo la figura 2, observamos que hay un falso positivo: el RTT estandarizado del primer hop supera el umbral dado por la τ del Test de Thompson. Podemos entender esto como una consecuencia de que el test utilizado no hace más que buscar valores atípicos (en nuestro caso atípicamente grandes) dentro de una muestra. Por lo tanto, en la medida que los RTTs entre hops no sean equitativos, no es de sorprender que casi siempre encontremos algún salto que sobresalga del resto. La clave está en que si tuviéramos un trayecto significativamente más largo, este valor que ahora resulta un outlier muy posiblemente quedaría opacado por el RTT diferencial de un verdadero salto continental.

Vale decir que en este caso particular, lo que parece estar sucediendo es que hay una cuestión técnica de la LAN desde la cual se dispara el *traceroute* que dificulta alcanzar el *gateway*, pues 80ms parece un tiempo elevado para esto. Independientemente, se probó llegar al mismo destino desde otra LAN con mejor tiempo de llegada al *gateway*, y sin embargo también se obtuvo un outlier en un hop posterior, reafirmando el punto anterior de que lo que se considera un outlier depende fuertemente de la escala de la ruta.

Universidad de Moscú

Recorrido en el Planisferio

A continuación se puede ver de forma bastante clara como el paquete tuvo que pasar por estados unidos para luego ir a Europa Occidental hasta llegar finalmente a Rusia.

RTT entre saltos

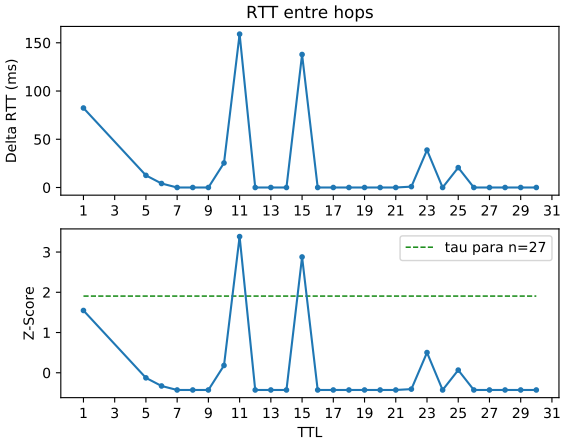

Figura 4: RTT entre saltos (antes y después de normalizar respectivamente) para el sitio `www.msu.com`. n es la cantidad de TTLs para los que se obtuvo un *time exceed* (que son los puntos que se grafican).

Para este caso, podemos observar en la figura 4, que se detectan dos saltos intercontinentales. Este caso tiene bastantes particularidades, en parte a la ruta compleja que se vió en el punto anterior.

El primer pico, que se da en el salto 11, coincide con un salto a un router de Roma, Italia (`89.221.41.171`). La cuestión es que el nodo desde el cual se realiza el hop también está ubicado en Italia (`185.70.203.32`), y dicho salto no fue detectado como outlier, pues en efecto tiene un RTT mucho más bajo. Una posible hipótesis que manejamos al respecto de porque ocurrió esto (que no se detecte el salto continental, pero sí el siguiente hop) es que el primer router al que se llega sea uno de los principales

gateways de Italia, lo que provoque que se encuentre congestionado y el paquete que se manda quede encolado un largo tiempo antes de forwardearse al siguiente hop.

El segundo pico, con TTL 15, presenta una situación similar: el destino está en Estados Unidos, pero los tres saltos anteriores también, y no fueron detectados. Acá hay una diferencia sin embargo: esos tres saltos mostraron todos RTTs diferenciales de 0, lo que implica que de hecho el RTT promedio de estos nodos fue menor que el promedio del último router que estaba en Italia. Este parece ser un caso de çamino asimétrico"(Jobst 2012). Razonablemente existe una mejor ruta desde Estados Unidos a Argentina, que cruzar toda Europa. Que la ruta no haya ido directamente por Estados Unidos, salteandose Italia, suena a una consecuencia de un cambio de estado de la red: por alguna razón el camino "óptimo"desde Italia a Rusia dejo de estar habilitado.

Universidad de Sidney

Recorrido en el Planisferio

A continuación se puede ver de forma bastante clara como el paquete tuvo que pasar por estados unidos para luego ir a Europa Occidental hasta llegar finalmente a Rusia.

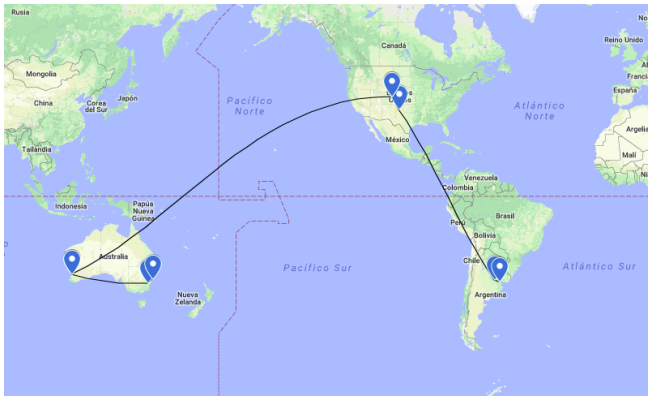

Figura 5: Recorrido realizado por los paquetes durante la ejecución de traceroute al intentar alcanzar el sitio `sydney.edu.au`

RTT entre saltos

A continuación intentamos ver los timepos entre saltos. Para esto calculamos la media de cada de los RTTs obtenidos de cada TTL para reducir a solo un valor las mediciones obtenidas por cada TTL.

Seguido de esto realizamos dos experimentos. Primero simplemente restamos los RTTs medios entre si como se puede ver en el primer gráfico, luego llevamos el experimento un paso mas allá normalizando con z-score los RTTs para eliminar cualquier tipo de deformación en las dimensiones del gráfico.

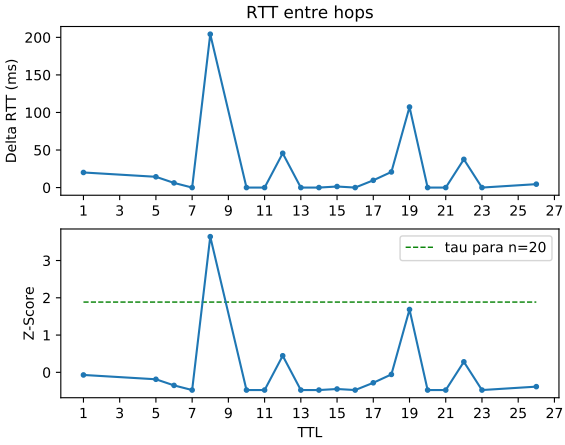

Figura 6: RTT entre saltos (antes y después de normalizar respectivamente) para el sitio `sydney.edu.au`. n es la cantidad de TTLs para los que se obtuvo un *time exceed* (que son los puntos que se grafican).

Como se puede observar ... COMPLETAR.

IV. Discusión

Como resultado final pudimos identificar de distintas maneras nodos destacados, por ejemplo a partir del grafo subyacente que armamos pudimos identificar en todos los casos al gateway a partir del nodo de mayor grado. A continuación describimos brevemente cada caso analizado por separado.

.1 São Paulo

En la red domiciliaria se pudo ver en el grafo resultante como destacaban dos nodos. Uno de ellos, el de mayor grado, era el gateway, el otro creemos que fue la computadora que tomó las mediciones ya que estaba accediendo a multiples sitios de internet lo cual provoco un gran intercambio de paquetes.

.2 Sidney

Este es el grafo mas grande y en el cual se vuelve aún mas claro el patrón que veniamos viendo, el nodo 10.2.203.254 resalta completamente de los demás y concuerda perfectamente con nuestra hipótesis de que el gateway es el que mas paquetes intercambia en este protocolo, lo cual lo vuelve un nodo destacado.

.3 Moscow

Quizas por una mala configuración de la red o por la poca clientela que había en este local en el horario en que se tomaron las mediciones (7:10 AM) es que este grafo resultó ser tan pobre, de todas maneras se pueden observar perfectamente los dos nodos que uno esperaría ver como minimo. El gateway y la computadora que tomo las mediciones.