Archipelago

- Santi Garcia Jimenez

Índice

- Introducción
- ¿Qué es?
- Medidas que estan tomando
- Validación de resultados
- Trabajo futuro
- ¿Donde encajamos nosotros?
- OFFTOPIC

Introducción

- Ark es un proyecto que deriva del skitter.
- Es llevado por el grupo CAIDA
- Colectivo formado por gente de los sectores comercial, gubernamental e investigación.
- Investigan aspectos teórico y prácticos:
  - Del comportamiento macroscópico de Internet, su infraestructura, comportamiento, uso y evolución.
  - Proveher un entorno colaborativo.
  - Mejorar la estabilidad de Internet.

Introducción

- Llevan desde 1999 haciendo medidas con skitter.
- Fueron los que inventaron el método mercator de identificación.
- Tienen datos en formato Pcap.
- Las Ips están anonimizadas(Misma máscara).
- Acceso a datos mediante acuerdo y estos no se pueden dejar a terceras personas.
- Hay acceso a determinadas medidas públicas.

¿Qué es?

- Ark es la evolución del skitter.
- Skiter tenia un monitor que hacía medidas.
- Ark va en un sentido mas distribuido con varias sondas que hacen medidas periódicas.
  - 27 monitores en 21 Paises.
  - En expansión.
  - Datos de
    - Rutas Ipv4 e Ipv6
    - Nombre DNS
    - Trafico DNS (respuesta y peticiones) trafico esnifado

Medidas que están tomando

- Trafico Ipv4
  - Tráfico paris traceroute ICMP hacia todas las rutas /24
  - Uso del SCAMPER
  - Agrupación de sondas para realizar el trabajo.
    - Grupos de 13 monitores encargados del traceo cada 48-56 horas
    - Por ciclo una única sonda hace las pruebas de alias.
  - Desde 17 septiembre del 2007
    - 1.3 billones de traceroutes 519 GB de archivos ”.WART”.

Medidas que están tomando

- Enlaces de AS
  - Mapas de ruta con RouteViews
  - Por cada grupo un archivo de rutas.
- Se hace un DNS Loockup de cada IP vista en las rutas.
  - Usan un volcado local para poder mirar con mas rapidez.
- TCPDUMP captura todas las peticiones DNS.
  - Se pueden bajar los datos de los ultimos 30 dias en formato PCAP.

Medidas que están tomando

- Pruebas de alias
  - Se usa una ventana de 20-24 ciclos con 2-3 dias por ciclo. (Unos dos meses de ventana para las Ips)
  - Usan
    - Iffinder (MERCATOR)
    - APAR (Metodo analitico de las /24)
    - Usan Ally la implementación de rocketfuel como validación (Solo usa UDP, hay que verificarlo)

Validación de resultados

- Adquisición de routers verdaderos de una red real y comparar.
  - Mendigar y suplicar por información (Debería de funcionar)
  - Listas de routers de WIDE y IIJ.(Hay que mirarlo)
- Validación contra Internet2
  - Validación de Ally y apar.
  - Yo aun no he encontrado planos de Ips. No se de donde los sacan.

Trabajo futuro

- Hacer mas fácil la creación y desplegado de experimentos.
- Capa de nivel superior para generación, captura y análisis.
- Permitir a terceros la creación de experimentos.
  - Separación entre usuarios y medidas (Orientan hacia Sandbox)
  - Inclusión de políticas:
    - Anchos de banda, selección de destinos, tipos de paquetes

Trabajo futuro

- Parece que estan orientando en hacer una especie de PlanetLab.
- Parece que quieren permitir la realización de experimentos por parte de usuarios.
- Estan trabajando sobre lenguajes para definición de experimentos(Basado en un sistema de tuplas llamado Marinda).
- Parece más un sistema de memoria distribuido o al menos de acceso distribuido.

¿Donde encajamos nosotros?

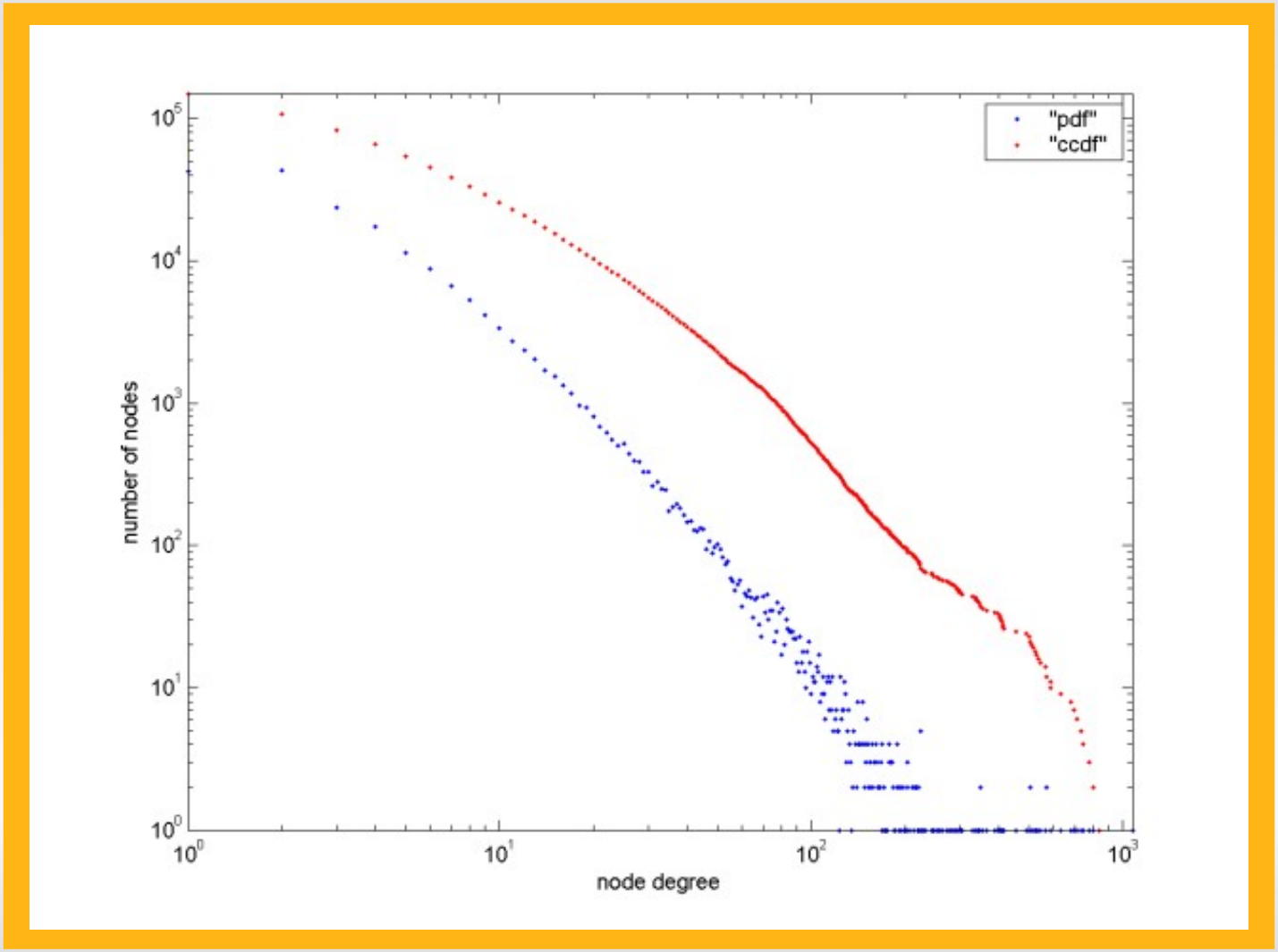
- Tenemos medidas que permiten con un 5% del total descubrir en torno a un 70% de los trues que se descubren por Ally(ICMP, UDP y TCP).
- No usan el metodo Ally por el problema de tener que hacer N^2 pruebas. Con esto lo podrían usar para conseguir una mejora en el alias y sin pruebas añadidas.
- Nuestro método se puede usar en un entorno distribuido donde el clustering por ttl no se puede usar.

OFFTOPIC

- En la presentación se dice que el router con mas aliases que se ha visto es uno de 100 interfaces. Particularmente, igual me parece muy grande.
- Hay una gráfica del número de interfaces en CAIDA que salen rangos de 1000 interfaces router. Esto ya me parece exagerado.

OFFTOPIC

- Gráfica:



OFFTOPIC

- Del formato de las trazas no he conseguido aun ver la especificación (WART). El programa de traducción de trazas es para freeBSD y no compila en un linux normal, al menos a las bravas.
- Las trazas anonimizadas en un principio no nos interesan, y de las otras ,en un principio, la licencia dice que no se pueden hacer pruebas hacia las Ips que saques de sus trazas.

OFFTOPIC

- Aún no me he conseguido explicar como usan el metodo APAR.
- Si no hacen pruebas origen-destino, controlando los dos puntos no tienen traceroute de ida y de vuelta.
- Supongo que harán una verificación con iffinder y una vez sacados aliases en dos puntos miran los traceroutes que pasan por ellos.
- Se reduce muchísimo la identificación de este modo de uso del sistema de identificación.