

INFORME PROYECTO ELO-322:
¿CÓMO FUNCIONA LA APLICACIÓN
SHAZAM?

MATÍAS DÍAZ

DANIEL MORALES

JAVIER ESCOBAR

PABLO REYES

PROFESOR AGUSTÍN GONZÁLEZ

28 de Julio de 2014

Resumen

¿Qué es Shazam?

Shazam es una aplicación que nos permite reconocer una canción que está sonando, con solo dejar sonar esta canción cerca de tu dispositivo móvil con la aplicación, este es capaz de identificar el título, autor y álbum de la canción.

¿Cómo funciona?

Al comenzar la aplicación esta creara una huella digital de audio y esta es comparada con la base de datos de millones de canciones, esta es enviada al servicio Shazam e inicia la búsqueda en la base de datos durante 10 segundos, si se encuentra devuelve la información al usuario, en caso contrario devuelve un error.

Para encontrar el esquema y su funcionamiento en la red se usaron herramientas enseñadas durante el ramo, posteriormente se hizo un análisis de los resultados entregados por estas herramientas con los cuales se pudo descifrar el funcionamiento a través de la red.

Introducción

Problema a resolver de Shazam

Imagina vas por la calle o estas en algún sitio cualquiera, y están escuchando la radio y escuchas una canción que te gusta y no has podido encontrar el nombre o autor de ninguna manera, la aplicación Shazam resuelve ese problema, ya que con shazam se puede reconocer cualquier canción y de forma sencilla en cualquier momento.

Reseña histórica

En el 2002 el servicio fue lanzado, y se conocía como '2580', los consumidores marcaban desde su teléfono este número para obtener el reconocimiento de la canción, el resultado te llegaba por mensaje de texto con el título y autor de la canción.

Con el tiempo y avance de las tecnologías Shazam fue agregando opciones para poder descargar las canciones que se identificaban, hasta llegar a lo de hoy, en donde aparece mayor información incluso como biografías, canciones recomendadas, conciertos, etc. Incluso ahora la aplicación no es solo utilizada para encontrar música, sino que también programas de televisión, series.

Limitaciones

Shazam puede identificar la música en cualquier lugar: desde la radio, la televisión, el cine o en una tienda, la limitación la encontramos en que sólo puede funcionar con música pregrabada, es decir, no puede encontrar canciones cantándolas o tarareándolas, de igual forma no podemos esperar que nos identifique ruidos.

Problema:

Shazam es una aplicación para celular que al poner el celular a "escuchar" una

canción que suena en el ambiente, este devuelve el nombre de la canción, el álbum en que aparece y el artista o grupo que la interpreta. Se quiere saber ¿Cómo funciona el streaming de música al servidor? ¿Cómo realiza el reconocimiento? ¿Cuáles son sus limitaciones técnicas?

Bosquejo de solución:

Se realizará una descripción de la aplicación, que servicios ofrece, su formato y plataforma, se estudiará su funcionamiento en la red, el protocolo para el streaming, la cantidad de servidores comprometidos, cómo mantiene sus servicios, la descripción de su operaciones y cómo lo hace para superar las dificultades y limitaciones técnicas tanto de su tecnología como la del dispositivo que hace la consulta.

Funcionamiento de Shazam

Para que Shazam pueda realizar la comparación de una muestra de audio de 10 segundos grabada desde nuestro teléfono con alguna canción de la base de datos de la aplicación de una manera rápida y efectiva no envía la muestra como un archivo de audio, sino que crea una huella de datos de audio a partir del espectrograma de frecuencias de la muestra de audio grabada. Para esto la aplicación genera a partir del espectrograma de frecuencias -que se muestra en la figura 1A, donde el eje de las X corresponde al tiempo, el eje Y corresponde frecuencia y la intensidad del color negro alrededor de distintos puntos del gráfico corresponde a la intensidad de alguna frecuencia de sonido en un tiempo determinado de nuestra muestra de audio- un espectrograma simplificado como se aprecia en la figura 1B, tomando los puntos de mayor intensidad del espectrograma original. A partir de esto la aplicación toma un punto de referencia o punto de anclaje de nuestro espectrograma simplificado y comienza a realizar mediciones hacia una zona objetivo, que consisten en almacenar en una tabla hash la diferencia temporal entre el punto de anclaje y algún punto de la zona objetivo, la frecuencia de ambos y el tiempo en el que comienza la medición, como se aprecia en la figura 1C y 1D.

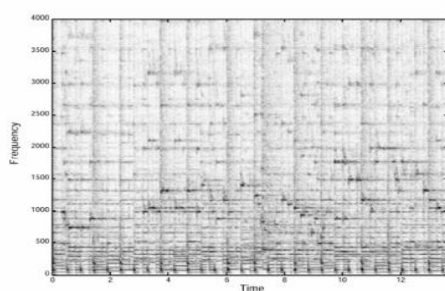


Fig. 1A - Spectrogram

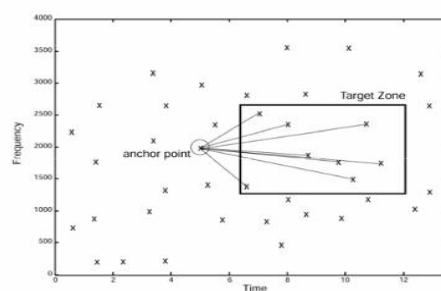


Fig. 1C - Combinatorial Hash Generation

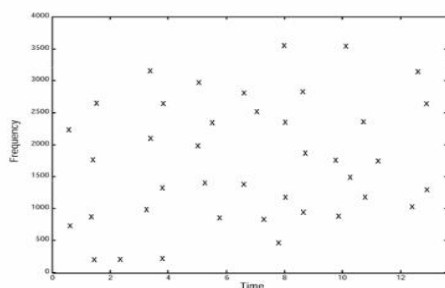


Fig. 1B - Constellation Map

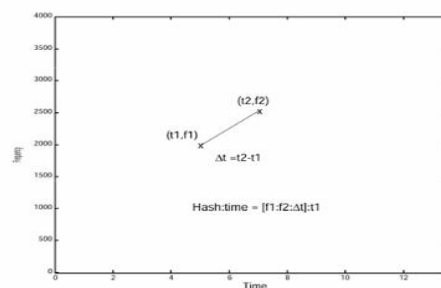


Fig. 1D - Hash details

Figura 1: proceso de creación de tabla hash y huella de datos

Luego estos datos son enviados a los servidores de Shazam donde se buscan las coincidencias de entre la tabla hash creada para nuestra muestra de audio y las canciones de la base de datos de la aplicación.

En las figuras siguientes es posible apreciar dos casos, el primero cuando la muestra de audio no coincide con una canción de la base de datos y el segundo caso cuando sí lo hace.

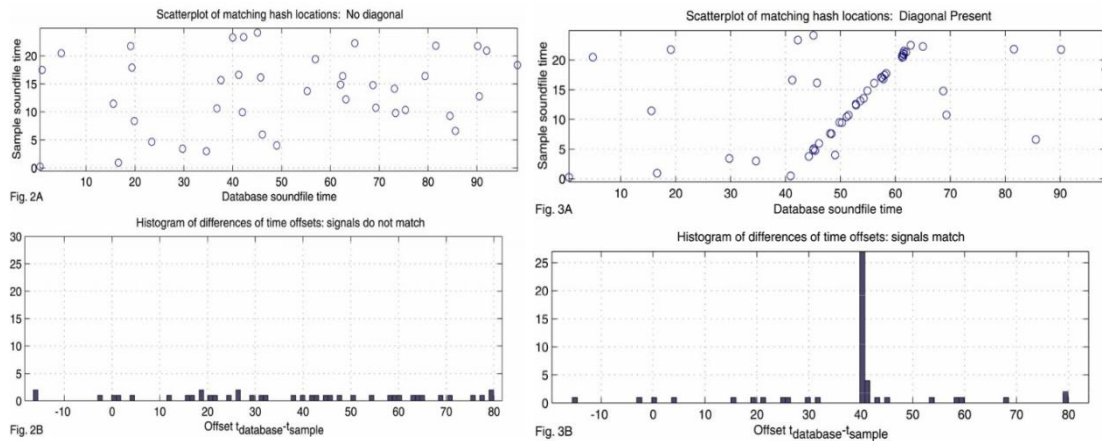


Figura 2 y 3: no coincidencia y coincidencia de audio respectivamente

En ambas figuras se aprecia que en los gráficos superiores el eje de las abscisas corresponde al tiempo de algún archivo de audio de la base de datos y el eje de las ordenadas corresponde al tiempo de nuestra muestra de audio. Los puntos son coincidencias de frecuencias entre ambos.

Cuando se encuentra una coincidencia de canciones debe ocurrir que se encuentre una sucesión de frecuencias coincidentes a medida que transcurre el tiempo en la muestra y en el archivo de la base de datos, así se formará una recta diagonal como se aprecia en la figura 3, y como es de esperar si las coincidencias de frecuencias entre la muestra y el archivo de la base de datos ocurren sin ninguna correlación temporal como en la figura 2, es claro que ambos archivos de sonido no son iguales por lo que la aplicación después de un tiempo te muestra en pantalla que la canción no está en la base de datos a menos que encuentre otra en la que suceda lo previo descrito en la figura 3.

Experimentación

Se procedió a encontrar la estructura a través de las redes del funcionamiento de la aplicación Shazam, para ello se dispuso de las siguientes herramientas:

Interceptor NG: Debido a que Shazam funciona en plataformas tipo Smartphone o tablets fue necesario usar una herramienta para "auditoriar" (sniff) las redes que fuera similar a Wireshark pero que funcione para este tipo de plataformas. Interceptor NG es una aplicación para Android que, si bien no es tan completa como Wireshark, nos permite capturar paquetes y entregar información suficiente para entender el tráfico de la aplicación.

NSlookup: Es una herramienta que informa sobre los dns autoritarios y no autoritarios de alguna dirección o dominio en particular.

Whois: es una herramienta que nos entrega una serie de información (nombres de dominio, direcciones, personas responsables, direcciones físicas entre otras) relacionada a una dirección o a un dominio.

El procedimiento fue el siguiente:

- 1) Se inicia la captura de paquetes a través de Interceptor NG.
- 2) Se pone en funcionamiento la aplicación Shazam.
- 3) Luego de terminar el análisis de la canción en Shazam se detiene la captura de datos.
- 4) Se revisa las IP con las que se conecta la aplicación a través de la herramienta "Whois" en una plataforma Linux (SO: Backtrack 5 r3).
- 5) De los resultados del análisis a través de Whois se busca en la web información sobre las empresas y nombres encontrados a través de "Whois".
- 6) Si en los resultados faltaba alguna empresa que especulábamos que intervenían en la red de Shazam la buscábamos a través de NSlookup para conocer su dirección IP y luego buscarla en el resultado de búsqueda de paquetes que se hizo con Interceptor NG.
- 7) Luego de identificar todas las funciones de las comunicaciones con las distintas direcciones IP mostradas por Interceptor NG se procedió a evaluar los resultados y crear un mapa de conexiones.

Resultados:

Las direcciones IP más recurrentes eran:

- 68.67.160.240 Que corresponde a la empresa AppNexus que se encarga de proveer de publicidad a las aplicaciones.

-209.15.194.216 Que corresponde a la empresa Peer1 que es una empresa proveedora de servidores y hosting.

-209.15.194.252 que corresponde también a Peer1, sin embargo a través de Nslookup y Whois esta dirección IP se encuentra inscrita por Shazam.com.

-5.44.16.252 Que corresponde a Shazam.com, sitio web de la aplicación.



Fig.3 Resultado de captura de paquetes a través de Interceptor NG donde se ve a 68.67.160.240 ,209.15.194.216 , 209.15.194.252 y 5.44.16.252



Fig.3 Resultado de NSlookup para Shazam.com

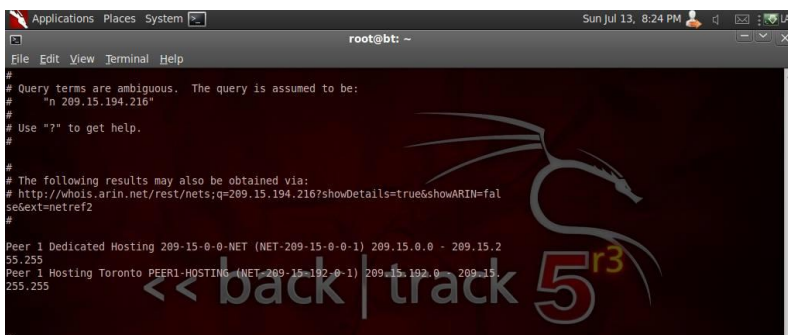
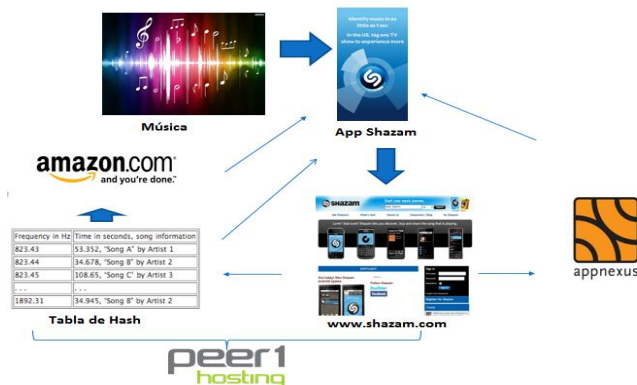


Fig.4 resultado de whois para la dirección 209.15.*.*

Conclusión a partir del Método Experimental

En respuesta a los resultados obtenidos en la parte anterior se pudo deducir y formar el siguiente esquema estructural con que se riges Shazam:



El cual se explica de la siguiente forma:

1°.- Cuando uno escucha una canción que está sonando en el ambiente y la quiere buscar, va a ejecutar la aplicación Shazam. Este a su vez va a crear las huellas de datos de audio a partir de espectrograma (obtenido desde la canción misma). Luego va a generar un hash a partir de mapa de constelación simplificando el espectrograma.

2°.- Luego la App Shazam se va a conectar y pasar el hash a la página www.shazam.com , con tal de que ella cree enlaces a distintas partes.

3°.- Uno de dichos enlaces ira a los servidores de shazam, más específicamente a su tabla de hash (cabe mencionar que tanto la página de shazam como sus servidores son arrendados a Peer1, que es a su vez uno de los proveedores de alojamiento virtual más grande del mundo, que cuenta con una extensa granja de servidores para arriendo).

La tabla de hash se encargara de comparar el hash entrante con los que miles que tiene ya almacenado, buscando así el hash que arroje el mayor número de coincidencias con el hash entrante. Cuando ya identificó el hash tira como salida la canción, autor y álbum del dicho hash.

4°.- Esta información va a parar directamente a la App Shazam como también a Amazon (compañía estadounidense de comercio electrónico y servicios de cloudcomputing a todos los niveles). El cual se encargará de ofrecer la compra y posterior descarga de la canción.

5°.- Volviendo a los enlaces que generó la página de Shazam, otro enlace que crea es hacia AppNexus (empresa con sede en New York que se encarga de ofrecer una plataforma de publicidad en línea en tiempo real), la cual se encarga de colocar los anuncios de publicidad y propaganda dentro de la App Shazam.

Con lo anteriormente dicho se explica más detalladamente cada paso que sigue la App Shazam (que se obtuvieron en la parte experimental), desde que se ejecuta hasta que devuelve la información de la canción en los distintos dispositivo móviles.

Referencias

- Documento escrito por desarrolladores de Shazam sobre su funcionamiento:
<http://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf>
- Código que explica su funcionamiento
www.royvanrijn.com/blog/2010/06/creating-shazam-in-java/
- Reseña histórica shazam
<http://es.wikipedia.org/wiki/Shazam>
- Funcionamiento general de shazam, limitaciones
<https://www.generacionyoung.com/tecnologia/mas-tecnologia/tecnologia-shazam-pasado-presente-futuro-app-reconoce-canciones/>
<http://www.shazam.com/>