Universidad del Valle de Guatemala

Algoritmos y Estructura de Datos

Jose Andrés Arenas #14470

Sebastian Arriola #11463

Fernando Figueroa #14175

Mayo, 2018

**Investigación Proyecto No. 2**

Un sistema de recomendaciones simplemente es una forma de poder analizar información de diferente manera y tener la capacidad de poder cruzar ese conocimiento con los usuarios para de una u otra manera saber sus gustos y así poder proveer recomendaciones en cualquier ámbito.

Aunque las personas no lo sepan directamente, el mundo vive en un constante uso de dichos sistemas, porque no siempre deben ser herramientas en línea o programas que nos digan explícitamente lo que queremos. De hecho, un ejemplo sencillo de un “sistema de recomendaciones” sería en el momento que le pedimos ayuda u opiniones a un familiar o amigo de algo. Esa persona se convertiría en nuestro sistema de recomendaciones, porque podría dar su opinión de lo que se le fue preguntado en base a la información y gustos que conoce de nosotros.

Ya a la hora de implementar ese tipo de situaciones en un programa inteligente si es algo más complejo que una persona. Para resolver este tipo de problemas, existen algoritmos capaces de hacer ese cruce de información mencionado al principio para poder dar soluciones acertadas a algún problema. Los algoritmos utilizados en estos sistemas son:

· *“Collaborative Filtering”*

· *“Content-Based Filtering”*

· *“Hybrid Recommendation Systems*

Mas adelante se definirán cada uno de ellos para comprender cómo cada algoritmo “ataca” el problema mencionado.

**Collaborative Filtering:**

Este algoritmo se basa en recolectar y analizar información de una cantidad muy grande de usuarios. La información del usuario que este algoritmo necesita es por ejemplo su comportamiento, costumbres, las actividades que realiza y preferencia en ciertos productos. Entonces el cruce de información es cuando a la hora que otro usuario que tenga similitudes con alguno otro usuario en la base de datos del algoritmo, este algoritmo predice que si al usuario A le gusta la bebida de tipo A, el usuario B que tiene similitudes en gustos, preferencias, etc, le gustaría también la bebida B entonces, se la recomienda. En otras palabras, solo es comparar gustos, preferencias y comportamientos entre muchos usuarios.

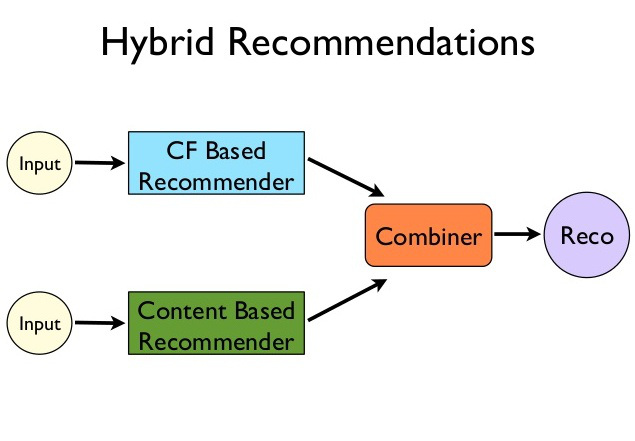
**Content Based Filtering:**

Este algoritmo se basa más en los productos que en los usuarios de los mismos, aunque sí utiliza las preferencias del usuario a la hora de hacer el cruce de información. En este sistema generalmente se usan palabras clave para definir el producto, y luego en base a estas palabras clave y las preferencias del usuario, se buscan otros productos que tengan características similares al que generalmente el usuario consume y esos son los que se le recomiendan. En otras palabras, se recomiendan los productos que más similitudes tengan entre todos los demás.

**Hybrid Recommendation System:**

Como su nombre lo indica, es un sistema que utiliza la unión de los dos algoritmos previamente descritos. Estudios han demostrado que es la mejor manera de resolver este tipo de problemas debido a que como utiliza dos algoritmos, los resultados tienden a ser mejores. Generalmente se usa uno, luego el otro y luego se unifica la información, o bien, haciendo un solo modelo que utilice alguna parte de las dos técnicas.

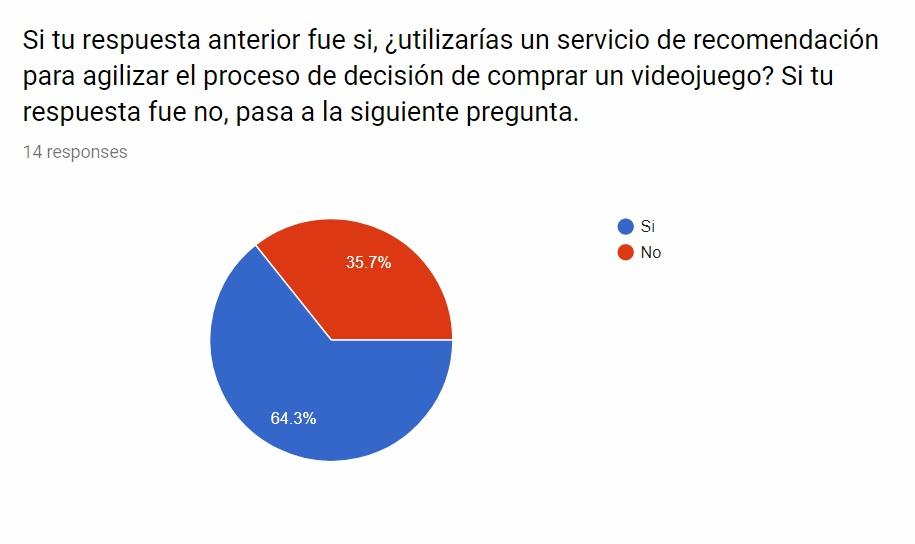
*Figura 1 - Sistema Híbrido de Recomendaciones*



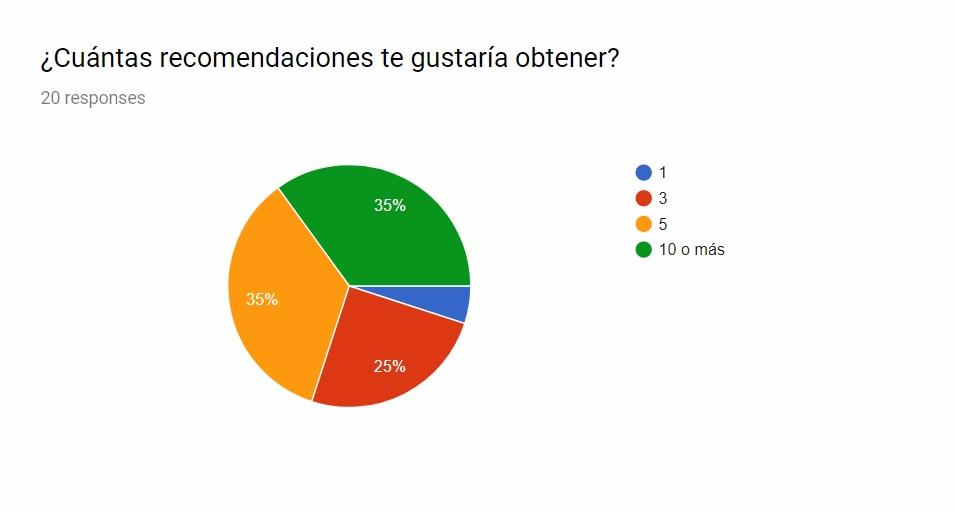
Design Thinking

* Fase 1 - Empatía - Encuesta: se realizó una encuesta a 20 personas por medio de Google. Los resultados más importantes se presentan a continuación:
  + En general, los usuarios estarían dispuestos a proveer información personal, como: nombre, edad, correo electrónico, gustos en videojuegos, cantidad de horas invertidas en jugar.
  + 64.3% de los encuestados les gustaría usar un servicio de recomendación para decidir qué videojuego comprar.
  + Un 90% de los usuarios quieren recomendaciones acerca de los videojuegos que sus amigos juegan.
  + Se mostró una inclinación a recibir desde 5 hasta 10 o más recomendaciones.

*Figura 2 - Encuesta, Pregunta 3*



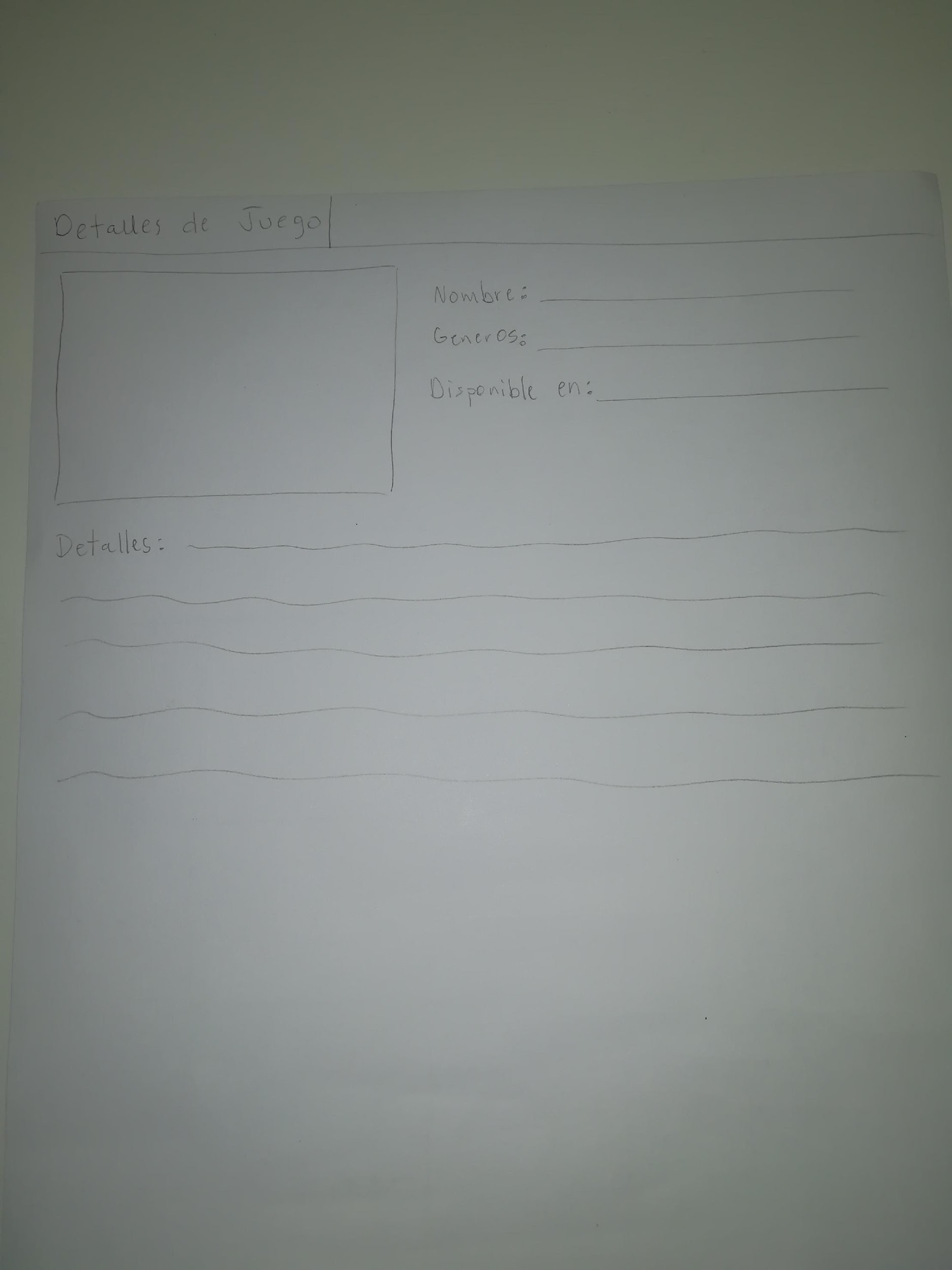
*Figura 3 - Encuesta, Pregunta 6*

**

*Figura 4 - Encuesta, Pregunta 7*

**

* Fase 2 - Definir el Problema
  + Con la información recabada en la fase de empatía, se observó que los gamers tienen interés en saber qué videojuegos están jugando sus amigos, el cual es un punto clave, que se puede atender por medio de un servicio de recomendación, en especial, uno basado en grafos, que permite relacionar amigos y videojuegos fácilmente.
  + Uno de los problemas de la comunidad Gamers es como llegar a una decisión acertada a la hora de adquirir nuevos videojuegos, debido a que en el mercado existen muchas opciones y la inversión en un juego es considerablemente alta. Por lo que se necesita alguna manera de poder facilitar dicha situación.
* Fase 3 - Lluvia de Ideas
  + Sistema que muestre los juegos más populares del mercado
  + Sistema de búsqueda a partir de palabras clave
  + Un sistema que funcione a partir de crear un perfil con algún tipo de información y en base a eso sugerir juegos.
  + Algún tipo de red social en donde el sistema sea capaz de hacer cruce de información con los gustos de los amigos que el usuario tenga y así sugerir juegos que tengan parecido a sus gustos.
  + Hacer nuevos amigos, que compartan gustos en videojuegos.
  + Los usuarios pueden hacer reseñas de los videojuegos, así como darles una puntuación.
* Fase 4 – Prototipos



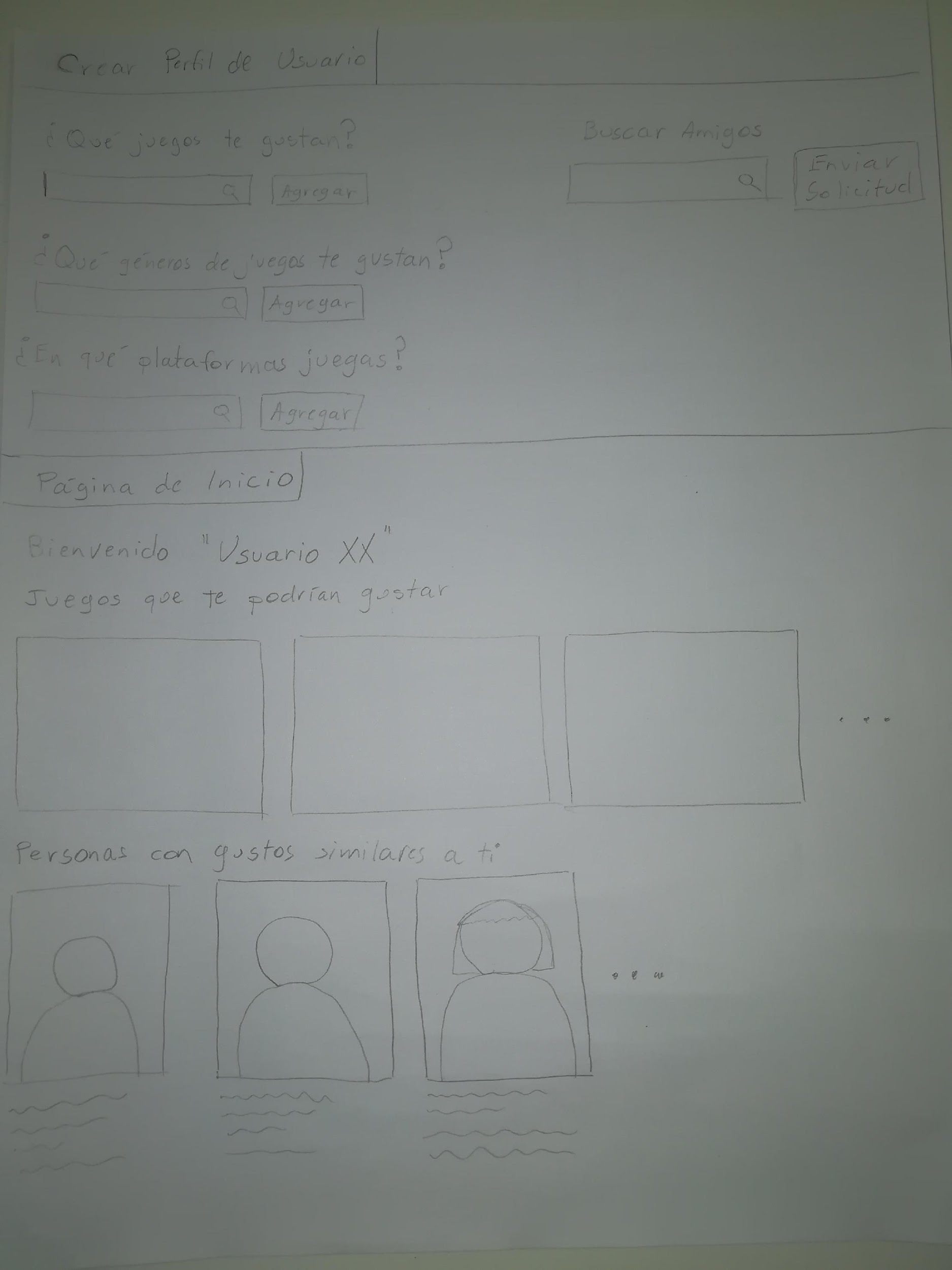


Foto #1 prototipo Foto#2 prototipo

* Fase 5 - Pruebas

El uso de nuestro programa para la recomendación de juegos resolvería el problema planteado de una manera eficaz de la siguiente manera. Como se puede ver en la imagen de Prototipo 1, a la hora que el usuario inicie el programa se le pedirá que ingrese una serie de datos para que se pueda crear su perfil. En ese momento, cuando termine de ingresar todos los datos requeridos, a la respectiva base de datos se le agregara un nuevo nodo que contenga al usuario.

Al tener el nuevo nodo creado en el grafo, se deben crear todas las relaciones con juegos, generos, plataformas, etc en relación a la información que el usuario haya proporcionado al sistema.

Con todas las relaciones del usuario ya creadas dentro del grafo de la base de datos, inicialmente ya se podría proceder a hacer algún tipo de cruce de información con los demás juegos y categorías ya existentes dentro de la base de datos.

A la hora que al sistema se le metan mas perfiles de otros usuarios con su respectiva información, la base de datos se haría mucho más grande y todo el sistema sería más funcional, pudiendo de esta forma resolver de una forma eficaz el problema definido dentro de la primera fase, porque existirían más relaciones entre toda la información del grafo, haciendo mejor el proceso de cruce de información para poder proveer las recomendaciones que el usuario solicite.

Diseño de Base de Datos y Modelado de Datos

Para el diseño de la base de datos, primero se decidió como modelar los datos. Para esto, se planteó una serie de preguntas, las cuales debían obtener sus respuestas de los datos.

1. ¿Qué videojuegos les gustan a mis amigos?
2. ¿Quiénes tienen gustos similares a mí?
3. ¿Qué videojuegos que no conozco podrían gustarme?
4. ¿Qué videojuegos son los más populares?

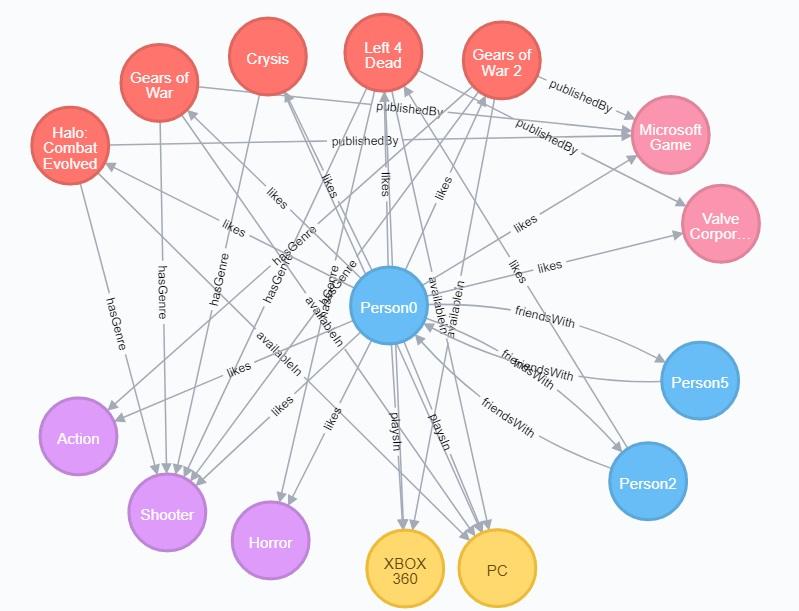
A partir de las preguntas, se identificaron las entidades que se debían modelar como vértices en el grafo, y las relaciones entre vértices. Se identificaron como sigue:

* Nodos
  + Personas
  + Videojuegos
  + Publicadores de videojuegos
  + Géneros de videojuegos
  + Plataformas de videojuegos
* Relaciones, se indican las entidades relacionadas utilizando una flecha, el nombre de la relación se indica entre brackets.
  + Amistad, (persona) --[friendsWith]--> (persona)
  + Juegos que me gustan, (persona) --[likes]--> (videojuego)
  + Géneros que me gustan, (persona)--[likes]-->(género)
  + Publicadores que me gustan, (persona)--[likes]-->(publicador)
  + Plataformas que utilizo, (persona)--[playsIn]-->(plataforma)
  + Género de videojuego, (videojuego)--[hasGenre]-->(género)
  + Plataforma de videojuego, (videojuego)--[availableIn]-->(plataforma)

Utilizando el modelo propuesto, se muestra en la figura 5 un nodo central que pertenece a un usuario, “Person0”, y todas sus diferentes relaciones con otras entidades:

* Person0 es amigo con Person2 y Person5
* A Person0 le gustan varios juegos: “Halo: Combat Evolved”, “Gears of War”, etc.
* A Person0 le gustan los géneros “Action”, “Shooter”, “Horror”
* Person0 juega en “PC” y en “XBOX 360”.
* A Person0 le gustan los publicadores “Microsoft Game Studios” y “Valve Corporation”.

*Figura 5 - Modelo de Base de Datos para un Usuario, Person0*



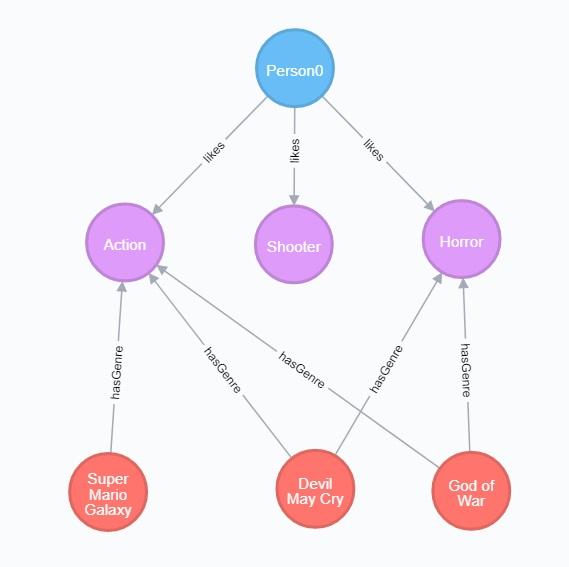
Algoritmos de Recomendaciones

Con el modelo de datos propuesto, se generaron distintos algoritmos de recomendación:

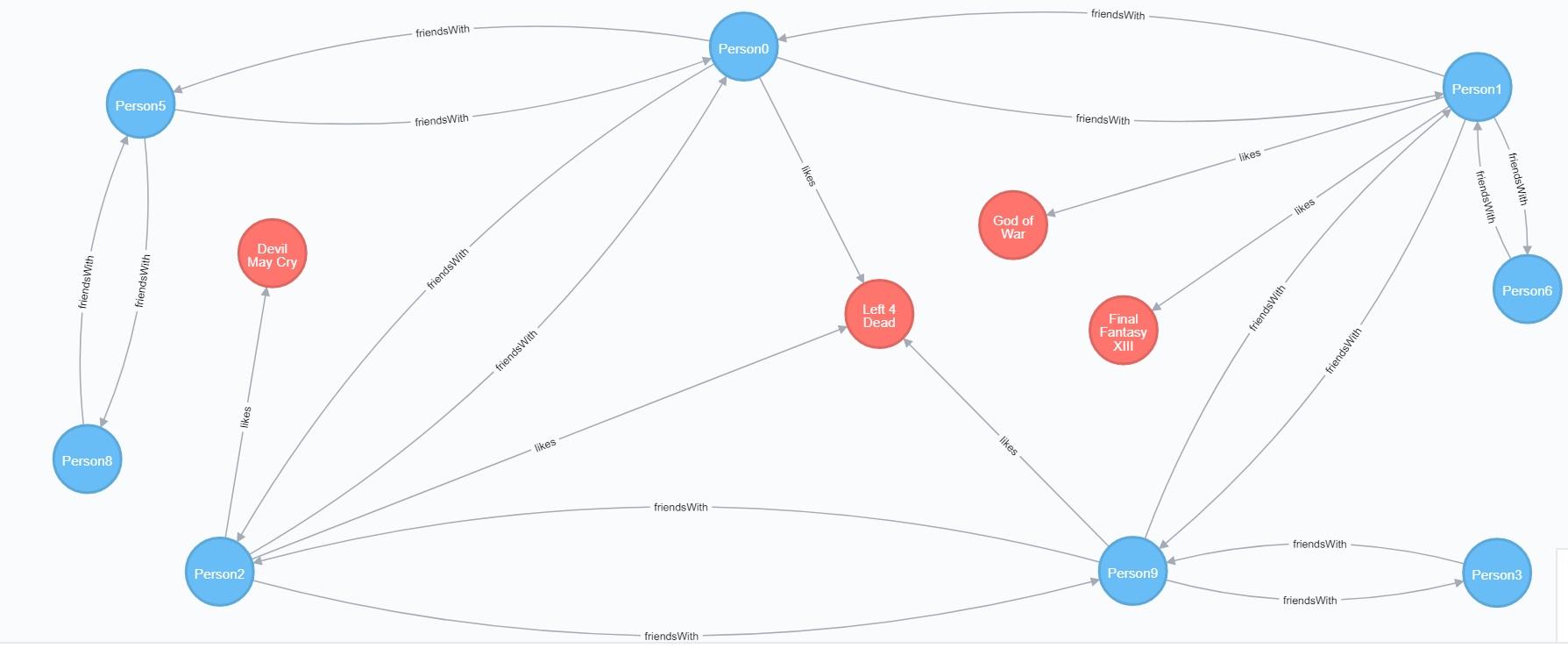
* Recomendación de juegos a un usuario, basado en los gustos del usuario.
* Recomendación de juegos a un usuario, basado en sus amistades.
* Recomendación de personas a un usuario, para generar nuevas amistades.

Se procede a explicar con detalle cada algoritmo de recomendación.

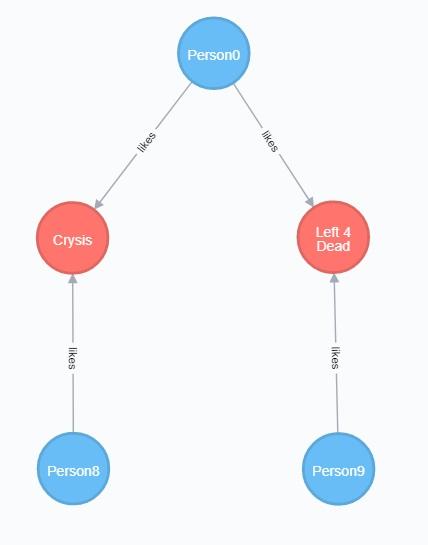
*Figura 6 - Recomendación por géneros para usuario “Person0”*



*Figura 7 - Recomendación por amistades para usuario “Person0”*

**

*Figura 8 - Recomendación de personas para usuario “Person0”*

**

1. Recomendación de juegos a un usuario, basado en los gustos del usuario. Ver figura 6.
   1. Los vértices morados representan los géneros que el usuario ha marcado como sus favoritos: Horror, Action y Shooter.
   2. Se realiza un recorrido para cada género, con el objetivo de encontrar todos los videojuegos asociados con dichos géneros.
   3. Los juegos que ya están asociados al usuario se eliminan de los resultados, para no recomendar un videojuego que que el usuario ya haya jugado.
   4. En los resultados, se observan los videojuegos: Super Mario Galaxy, Devil May Cry 4, y God of War.
   5. Se le da prioridad a los resultados que tengan más géneros asociados, más bien, a los videojuegos que aparezcan con más frecuencia.
   6. El resultado de esta búsqueda serían los videojuegos Devil May Cry 4 y God of War. Super Mario Galaxy quedaría excluído, basándonos en el criterio de la frecuencia de aparición.
2. Recomendación de juegos a un usuario, basado en sus amistades. Ver figura 7.
   1. Los vértices azules representan personas. Las recomendaciones generadas son para el usuario “Person0”.
   2. Se realiza un recorrido breadth-first, únicamente en las aristas marcadas como “friends With”. De esta manera, se obtiene los amigos del usuario “Person0”, que son: “Person1”, “Person2”, y “Person5”.
   3. Para cada amigo, se realiza un recorrido breadth-first, únicamente en las aristas marcadas como “likes” que asocien a un usuario con un videojuego.
   4. Se obtienen los videojuegos: Devil May Cry 4, God of War, Left 4 Dead, y Final Fantasy XIII.
   5. Debido a que Left 4 Dead también está asociado a el usuario “Person0”, se elimina de los resultados.
   6. El resultado de la búsqueda sería entonces: Devil May Cry 4, God of War, y Final Fantasy XIII.
3. Recomendación de personas a un usuario, para generar nuevas amistades. Ver figura 8.
   1. Los vértices azules representan personas. Las recomendaciones generadas son para el usuario “Person0”.
   2. Se realiza un recorrido breadth-first, únicamente en las aristas marcadas como “likes” que asocian a “Person0” con videojuegos.
   3. Para cada videojuego, se buscan todos los vértices marcados como personas, que tengan un arista a dicho videojuego marcada como “likes”.
   4. Se eliminan de los resultados aquellas personas que ya sean amigos con el usuario “Person0”.
   5. Los resultados de la búsqueda son entonces “Person8”, y “Person9”.