Master Degree Project

UNIVERSITY OF SKÖVDE

RECOMMENDING STEAM GAMES TO USERS

Matching games based on the time profile of a user

Master Degree Project in Informatics with a Specialization in Data Science
One year level, 15 ECTS
Spring term 2019

Niclas Elmäng

Supervisor: Jiong Sun
Examiner: Jiong Sun

1 Introduction

Steam today has a discovery queue that recommends game titles to users. This queue is made up of new releases, popular titles and games recommended to a user based on a lot of factors. Some users can spend a lot of hours every week to play certain games while other users maybe only have time to spend a couple hours, or even less, per week. One area that the current recommender system might fail at is finding the most relevant game based on how much time needs to be spent on it. Not a lot of research has been done on this type of recommender system. Users with a limited time for playing games might appreciate being able to complete games in a reasonable amount of time, instead of spreading one game session over weeks if not months. The same can be done for those with much more time available. With the amount of indie games available there might to be some game/hidden gem that satisfy the genre and range of time necessary to suit the unique experience per each user.

2 Background

2.1 Recommender system

Recommender systems have gained popularity both in online applications such as e-commerce, or video-on-demand (streaming) services, and research applications. They are used to, like the name suggests, recommend products to users based on some set of metrics, (Shani & Gunawardana, 2008). An e-commerce store might use such a system to show products that are frequently bought together, while a video streaming service might use it to pair videos to users based on how similar the user's interests are.

2.1.1 Collaborative filtering

Collaborative filtering is the process of predicting what would be of interest based on many other neighboring data points. For example, a user-based collaborative filtering uses many previous user's information about a product. One use case is predicting movie ratings where the rating done by a user is predicted by looking at other users who have rated the same movies similarly. (Balabanovic & Shoham, 1997)

Item-based collaborative filtering is another type of filtering with the difference being what is compared. In this type of filtering the items are compared to each other. For example, given the previous data showing that some products are usually bought together. Using this kind of filtering the user can get recommended other items that they otherwise might not have considered, since those types of products often are bought in conjunction. (Balabanovic & Shoham, 1997)

2.1.2 Content-based

Content based filtering will try to pick items that are similar to what a user already likes. This is done by comparing the features of the items, e.g. in movies the actors, director and title. This works without using other users since only the movies, in this case, are compared to each other and the user is given other movies with similar features. (Balabanovic & Shoham, 1997)

2.1.3 One Hot Encoding

In natural language processing, one hot encoding is the process of turning categorical words into vectors of numerical value. This is needed to be able to do calculations on the categories, such as similarity calculations or for use in machine learning.

2.2 Evaluation

Offline experiments, as mentioned by Shani & Gunawardana (2008) is the process of using pre-collected data and simulate the behavior of the users. This kind of experiment has the advantage that no user-interaction is needed, and the precision and recall can be calculated by hiding some of the user data for testing. This kind of evaluation relies on having some ground truth that can be compared against. The balance between precision and recall lies in how many recommendations are given. More recommendations are likely to increase the recall while reducing the precision. In the case of a recommender system where the number of recommendations is set ahead of time, the precision is the most interesting measure.

User studies, (Shani & Gunawardana, 2008) is another experiment that has the advantage of being able to answer many different questions. This relies on gathering users and letting them test the recommender system. In this case the behavior when using the system can be tested. The main disadvantage with user studies is the fact that it is costly to arrange, both in terms of time and cost, if the testers are paid.

2.3 Related Work

Sifa, Bauckhage & Drachen (2014) implemented a game recommender system using archetypal analysis [5] with the intent to make it easier for users to find relevant games fit to their specific interests. They base this on the fact that there already exists a huge number of games and it is increasing quickly. Their recommender model uses implicit feedback by hiding some of the games in an attempt to recommend it back to the user. Their model looks at how much time the user spent on a game and the preferences the user has towards that kind of game based on the genre. Using precision and recall for evaluating their performance they managed to get successful results when using the archetypal analysis model.

O'Neill et.al. (2016) performed an extensive data gathering operation on the Steam platform, gathering data on all info on the games, what games all users have played and what the users friend network looked like. This study was more of as exploratory analysis on the Steam network; thus, it does not implement any kind of recommender system. They presented some interesting data like how diverse the games are in user friend groups or how many games are owned or played by users. They mention that the top 20% Steam users are responsible for the top 73% total market value of owned games, which is an interesting fact about the userbase. One of their contributions was to release their dataset, consisting of everything there is on the Steam network, for the public to use.

3 Problem definition

Today users have an increasing difficulty of finding games that may be of interest to them. With the sheer amount of games that gets published every month on distribution platforms such as iOS App Store (PocketGamer.biz, 2019) and Steam (“Number of games”, 2018), the process of searching for, and finding, those games that are most fit to the gamer is bound to be difficult, if done by the user. Some users have different demands or constraints than others, e.g. they only have a set number of hours available to play games during a short span of time, or they only like a certain kind of genre.

Games vary a lot in how much time usually is spent on them, as can be seen when using SteamSpy’s service (“Top By Playtime”). Being able to complete a game in a reasonable amount of time, with respect to how much time the user usually spends in two weeks, can be something many users would appreciate. A recommender system that helps users find games with similar genres and a more fitting required playtime is proposed.

A problem that the related work does not consider is the users playtime behavior. They only consider how much time a user has spent on a specific game/games and then find recommendations based on that. What is proposed here is to make a recommender system that tries to achieve some serendipity (unexpected positive results), as well as similarity, as mentioned in (Shani & Gunawardana, 2008).

3.1 Method

A recommender system using content-based filtering with extra weight added to the user’s personal time profile and preferred genres has been implemented. The reason content-based filtering is used is because the recommendations are supposed to find interesting and fitting games for the specific user based on their own playtime behavior and genre preferences. This information is retrieved implicitly from the dataset as the ground truth.

A dataset consisting of steam users, their games played and play times, and some metric showing time spent per week was used to create this recommender system. The dataset, as mentioned before, comes from O’Neill et.al. (2016) and some pre-processing was needed to get a good amount of relevant data collected.

The performance of the recommender system should be evaluated by using a user study, but the time frame of this project would not allow that. A precision/recall measure is unable to be evaluated in this type of recommender system because there is no good ground truth to compare the recommendations against.

4 Implementation

4.1 Data Mining

The dataset from O'Neill et.al. (2016) is about 170 GB in size. This was reduced to about 80 MB of information about some users and their games. The user's playtime behavior was calculated as the playtime in total over two weeks, while the most preferred genre was retrieved from the user's most played game in that period. These features are what the recommendations are based on.

A lot of data has been discarded, only a fraction of the actual games dataset (about 50GB in size) remains. The dataset is sampled with a very low fraction which is not preferable to just picking a pre-determined number of users and retrieving their game information. This was not feasible on the current hardware; thus the data was randomly sampled and might bring up some problems where enough data about each specific user is missing.

4.2 Feature Selection

To give good recommendations with the user's periodical playtime behavior in mind the time they spent on playing games was selected as a feature, which was retrieved implicitly as mentioned in 4.1. To also recommend similar games the genre of the game was selected as another feature. This feature had to be converted from text to numbers, done by one hot encoding, to be used in the similarity function. The game's release date was also used to make the results more constrained and increase serendipity, i.e. unexpected positive outcome.

4.3 Similarity Function

The process of finding games with the constraints of the user's playtime behavior and their genre preferences was done by calculating a weighted cosine similarity on the features. The genres are weighted with a value of 1.5 while the playtime was weighted using 1.0. The release date was weighted lower using a value of 0.5.

An un-weighted similarity function was attempted but the results were way too diverse. The constraints the release date feature came with meant that a lot of similarly games with respect to genre were not considered. This led to the process of implementing a weighted system instead, where the genres and playtime are weighted higher than the release date.

5 Evaluation

As mentioned before, the evaluation could not be conducted by calculating precision and recall. This was because of the nature of the types of games the recommender system is supposed to recommend. Since it is supposed to give games with a similar total playtime, genre and, to an extent, the same release year, it will not give back exactly the games the user already has played. This is because the user must have spent the average amount of time anyone spends on the game in forever, just during the two weeks that the data consists of. The user needs to be an average person since forever in just a mere two weeks. This fact rules out the testing method that removes a game from the user's list of preferred games in an attempt to recommend it back to the user, like Sifa et.al (2014) did in their study.

Subjective analysis of the recommendations and what they are based on show relatively satisfying results, see Figure 1. These recommendations are much more diverse, which could increase serendipity based on how the user would choose to react. Since the performance can only be measured by having pre-existing knowledge about the recommended games, this type of recommender system should be evaluated by conducting a user study to get the user's opinions. While conducting the user study the actual preference of the users can be established, such as how much time is spent each week or what genres are the most interesting.

```
Based on a playtime of 2223 minutes by user: 76561198061281162 and their recently most played game:
```

<table>
<tr><th></th><th>appid</th><th>Title</th><th>Release_Date</th><th>negative</th><th>positive</th><th>developer</th><th>publisher</th><th>Rating</th><th>Genre</th></tr>
<tr><td>0</td><td>201790</td><td>Orcs Must Die! 2</td><td>2012</td><td>685</td><td>13330</td><td>Robot Entertainment</td><td>Robot Entertainment</td><td>83</td><td>Action|Indie|Strategy</td></tr>
</table>

```
Recommended games with similar playtime and genre, sorted by recommended:
```

<table>
<tr><th></th><th>appid</th><th>Title</th><th>Release_Date</th><th>negative</th><th>positive</th><th>developer</th><th>publisher</th><th>Rating</th><th>Genre</th></tr>
<tr><td>0</td><td>233450</td><td>Prison Architect</td><td>2015</td><td>2306</td><td>31751</td><td>Introversion Software</td><td>Paradox Interactive</td><td>83</td><td>Indie|Simulation|Strategy</td></tr>
<tr><td>1</td><td>221380</td><td>Age of Empires II HD</td><td>2013</td><td>3188</td><td>54222</td><td>Skybox Labs, Hidden Path Entertainment, Ensemb...</td><td>Microsoft Studios</td><td>68</td><td>Strategy</td></tr>
<tr><td>2</td><td>3900</td><td>Sid Meier's Civilization® IV</td><td>2006</td><td>129</td><td>1400</td><td>Firaxis Games</td><td>2K</td><td>94</td><td>Strategy</td></tr>
<tr><td>3</td><td>238010</td><td>Deus Ex: Human Revolution - Director's Cut</td><td>2013</td><td>1236</td><td>13071</td><td>Eidos Montreal, Feral Interactive (Mac)</td><td>Square Enix, Feral Interactive (Mac)</td><td>91</td><td>Action|RPG</td></tr>
<tr><td>4</td><td>48240</td><td>Anno 2070™</td><td>2011</td><td>3691</td><td>5658</td><td>Blue Byte, Related Designs</td><td>Ubisoft</td><td>83</td><td>Strategy</td></tr>
<tr><td>5</td><td>2820</td><td>X3: Terran Conflict</td><td>2008</td><td>272</td><td>1637</td><td>Egosoft</td><td>Egosoft</td><td>73</td><td>Action|Simulation|Strategy</td></tr>
<tr><td>6</td><td>208480</td><td>Assassin's Creed® III</td><td>2012</td><td>3179</td><td>9377</td><td>Ubisoft Montreal</td><td>Ubisoft</td><td>80</td><td>Adventure</td></tr>
<tr><td>7</td><td>22490</td><td>Fallout: New Vegas</td><td>2010</td><td>2153</td><td>46753</td><td>Obsidian Entertainment</td><td>Bethesda Softworks</td><td>84</td><td>Action|RPG</td></tr>
<tr><td>8</td><td>49520</td><td>Borderlands 2</td><td>2012</td><td>11162</td><td>145039</td><td>Gearbox Software, Aspyr (Mac), Aspyr (Linux)</td><td>2K, Aspyr (Mac), Aspyr (Linux)</td><td>89</td><td>Action|RPG</td></tr>
<tr><td>9</td><td>91200</td><td>Anomaly: Warzone Earth</td><td>2011</td><td>481</td><td>3299</td><td>11 bit studios</td><td>11 bit studios</td><td>80</td><td>Action|Indie|Strategy</td></tr>
</table>

```
13 out of 19 genres match
```

Figur 1 Games recommended based on a playtime, genres and the release date.

The first attempt of similarity calculation was done using only the playtime and the genre as features. This very limited 2-dimensional feature set led to a lot of similarities of 1.0, as seen in Figure 2. This is because there is a sufficient amount of games with varying average playtime such that any given user's weekly playtime is bound to be exactly like at least some game's playtime, while also matching on the genres.

```
Based on user: 76561197994503872

and their recently most played game:
```

	appid	Title	Genre
0	8880	Freedom Force	RPG|Strategy

```
Recommendations based on user's recent playtime and game:
```

	similarity	appid	Title	Genre
0	1.000000	228280	Baldur's Gate: Enhanced Edition	RPG|Strategy
1	1.000000	33670	Disciples III - Renaissance Steam Special Edition	RPG|Strategy
2	1.000000	61700	Might & Magic: Clash of Heroes	RPG|Strategy
3	1.000000	3170	King's Bounty: Armored Princess	RPG|Strategy
4	0.999999	61520	Age of Wonders Shadow Magic	RPG|Strategy
5	0.999999	202710	Demigod	Action|RPG|Strategy
6	0.999999	57740	Jagged Alliance - Back in Action	RPG|Simulation|Strategy
7	0.999999	237890	Agarest: Generations of War	Adventure|RPG|Strategy
8	0.999999	208400	Avernum: Escape From the Pit	Indie|RPG|Strategy
9	0.999999	48220	Might & Magic: Heroes VI	RPG|Strategy

Figur 2 Early version of the recommender system with only playtime and genre as similarity feature.

6 Conclusion

6.1 Discussion

The problem of recommending games, other than just taking the most popular, highest rated one now, is an interesting topic. Variations of it already exists, Sifa et.al. (2014). Accommodating a certain type of constraint in the users is something relatively unheard of, at least in the case of giving them an appropriate gaming experience in terms of time to complete. The, albeit subjective, results here show that such a game recommender system could be made if enough features are compared and tweaked and a better evaluation method was conducted. The problem with the current implementation is that some games, even though they have the same genre and similar playtime, can be vastly different in who they are for. Someone playing a game with a recommended age of 18+ could suddenly be recommended a game for ages 7+. Though this phenomenon could lead to increased serendipity, as talked about before, it should be adjustable to fit the user's preferences.

6.2 Future Work

A clear improvement needed is an actual user study, which was not possible in time frame of this project. This would give quantifiable results that could be used to improve this model, as well as being compared to the results of other studies. Depending on how the user study is conducted even more questions about user behavior could be answered and used for improvement.

The model could be improved by bringing the developer and publisher into the calculation. This would be good if the user likes the types of games a certain developer makes, but the weights in the similarity function might need adjusting. As mentioned in 6.1, the age restriction for a game might be beneficial to implement. Finding games that fit the time aspect, genre and the age restriction might lead to the biggest increase in evaluation quality.

Instead of using a discrete playtime value, the time could made into an interval instead. This would increase the diversity of the recommendations while still staying true to the time behavior and genre preference of the user. Using this interval, it would be beneficial to allow the users to choose their own preferred playtime and genre, maybe even select a game to base some other features on like developer, director, artists or title.

The recommender system could also be allowed to recommend games with genres the user rarely play to give them the opportunity to discover something completely new that they never would have considered before, as is the purpose of the recommender system. This might lead to another increase in serendipity.

7 References

Balabanovic, M., Shoham, Y. (1997) *Content-Based, Collaborative Recommendation.* Available online: https://www.researchgate.net/profile/Marko_Balabanovic/publication/220426798_Fab_Content-Based_Collaborative_Recommendation/links/55af6ec208ae6aa568b3a7ef.pdf [2019-05-14]

Number of games released on Steam 2018 | Statistic. (2019). Available online: https://www.statista.com/statistics/552623/number-games-released-steam/ [2019-04-09]

Number of Steam users 2018 | Statistic. (2018). Available online: https://www.statista.com/statistics/308330/number-stream-users/ [2019-04-09]

O'Neill, M., Vaziripour, E., Wu, J., Zappala, D. (2016*) Condensing Steam: Distilling the Diversity of Gamer Behavior.* Available online: https://steam.internet.byu.edu/oneill-condensing-steam.pdf [2019-04-20]

PocketGamer.biz: App Store Metrics (2019), https://www.pocketgamer.biz/metrics/app-store/ [2019-05-14]

Shani & Gunawardana (2008) *Evaluating Recommender Systems*. Fourth International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution. DOI: 10.1109/AXMEDIS.2008.21 [2019-05-14]

Sifa, Bauckhage & Drachen (2014), *Archetypal Game Recommender System*. Available online: https://pdfs.semanticscholar.org/f057/a249c97f908d8e262cffab709c93be9d2314.pdf [2019-04-09]

Top By Playtime. (n.d.). SteamSpy - All the data about Steam games. Available online: https://steamspy.com/ [2019-04-09]