Attainment Ratings for Graph-Query Recommendation

Hal Cooper
Columbia University
500 W 120th Street
New York, New York, USA 10027
hal.cooper@columbia.edu

Garud Iyengar
Columbia University
500 W 120th Street
New York, New York, USA 10027
garud@ieor.columbia.edu

Ching-Yung Lin
Graphen, Inc.
500 5th Avenue, Floor 46
New York, New York, USA 10110
cylin@graphen.ai

ABSTRACT
The video game industry is larger than both the film and music industries combined. Recommender systems for video games have received relatively scant academic attention, despite the uniqueness of the medium and its data. In this paper, we introduce a graph-based recommender system that makes use of interactivity, arguably the most significant feature of video gaming. We show that the use of implicit data that tracks user-game interactions and levels of attainment (e.g. Sony Playstation Trophies, Microsoft Xbox Achievements) has high predictive value when making recommendations. Furthermore, we argue that the characteristics of the video gaming hobby (low cost, high duration, socially relevant) make clear the necessity of personalized, individual recommendations that can incorporate social networking information. We demonstrate the natural suitability of graph-query based recommendation for this purpose.

CCS CONCEPTS
-Information systems → Graph-based database models; Recommender systems; -Human-centered computing → Social recommendation;

KEYWORDS
ACM proceedings, graph-based database models, recommender systems, social recommendation

ACM Reference format:
Hal Cooper, Garud Iyengar, and Ching-Yung Lin. 2018. Attainment Ratings for Graph-Query Recommendation. In Proceedings of Twelfth ACM Conference on Recommender Systems, Vancouver, Canada, October 2018 (RecSys’18). 5 pages.
DOI: 10.475/123_4

1 INTRODUCTION
Steam [2] is the largest digital distribution service for video games on the PC, Mac, and Linux platforms, with more than 18,000 available video game products. Products on Steam are actively marketed to (more than 100 million) users, with regular discounted sales events and pop-ups of recommended products.

Video distributions services such as Netflix [12] and Amazon [22] have long used recommender systems to help customers become aware of suggested products. Netflix is the runner of the well-known “Netflix Prize” [25] for recommendation systems using user-movie rating triplets. These services primarily employ “explicit” ratings data, wherein a user directly and deliberately inputs some form of rating for a product. In this context, implicit data, i.e. data that may implicitly indicate user preferences, but that does not involve the user explicitly designating a score or writing a review, is often seen as less causally informative, more biased, and less precise. Unary ratings (where it is possible to “like” a product e.g. click on an ad, but not possible to register dislike), are common in implicit rating scenarios [16] and exemplify many of the aforementioned issues. Though Steam does allow users to post reviews of purchased products, the use of such data in making recommendations could be called into question due to review bombing [20]. Fortunately, there exists a wealth of additional data less susceptible to manipulation.

Steam is also a software platform on which purchased products are managed, organized, and run. It is the interactive nature of video games that distinguishes games from other forms of media. Rather than a passive process of consumption, playing video games involves making conscious choices about how, when, and why to perform certain interactions. It is this interactive experience of the consumers with the product that they purchase (and other users who have made the same purchase) that differentiates digital distribution systems like Steam from retailers like Netflix and Amazon that sell physical and digital products through an online marketplace. Furthermore, in addition to being a marketplace, Steam is both an active online community with a well-defined social network. In this paper, we argue that the choices players’ make are indicative of their preference of titles, and can therefore be used for the purposes of product recommendation. Furthermore, we argue that presently recorded data directly reflects these actions.

The boom in multiplayer online video games occurred around the same time as the increase in the popularity of online social networks [9]. Perhaps in response, Microsoft introduced the Xbox Live system [1]. In addition to centralizing and streamlining online play, Xbox Live was created as a social network. Users had avatars that represented their likenesses, and added other users, known in real-life, or encountered during online play, to their list of friends through the use of unique “GamerTags”. The concept of an “Xbox Live Achievement” (hereafter simply referred to as an “achievement”) was created, “gamifying” the play of video games [18]. Achievements are awarded when players completed certain (potentially difficult) tasks in a game, and these then contribute to a “GamerScore”. This GamerScore, as well as the achievements comprising it, would be publicly visible, giving users a form of bragging rights that demonstrated their skill. This “gamification” of playing video games (that is, giving the process of playing and completing video games the property of a game at a meta-level) served to entice...
players to become more competitive (therby participating in more play and being more invested in the Xbox Live social network) and makeaditional purchases (with so called “play-to-win” models now being a common method for monetization of game products [24]).

Digital marketplaces for other consoles [26] and computers (e.g. Steam) soon followed suit. Through the use of the Steam launcher or website, Steam members can view the products owned by other members, as well as their Steam Achievements. Users can also view the global stats of achievements, such as the percentage of owners of a game who have gained a particular achievement, as well as the requirements for gaining it (though some particularly difficult achievements, or achievements that spoil the game, may have their details hidden). This data is also available through the Steam WebAPI [29]. Prior to April 10, 2018, the vast majority of this data was publicly accessible. After this date, most of the data was made private [8] through the retroactive replacement of a user opt-out procedure with an opt-in procedure.

2 ATTAINMENT RATINGS

In this paper we argue that achievements are highly informative implicit data that can be used to determine a score representing how much a particular user likes a particular game. The rationale is quite simple; the more a user enjoys a game, the more they will attempt to complete everything it has to offer. We compute an overall achievement score for a particular user by combining each individual achievement weighted by difficulty level computed using global achievement rates.

Hours played has been proposed in the literature [14] as a metric to approximate user preferences. However, there are significant issues with the use of simple play times. For instance, the length of games is not uniform. Certain genres (e.g. Role-Playing games) are typically significantly longer in length than other genres (such as Action games), and many highly acclaimed games are extremely short (such as the Playstation title, Journey).

We propose to eliminate these shortcomings by using achievement data to define what we call an attainment rating. The attainment rating is computed by combining each individual achievement weighted by a difficulty level computed using global achievement rates.

From (2) it follows that that attainment rating makes a large contribution to $A_{gs}$. Similarly, extremely common achievements, i.e. those with $C_{gi} \approx 1$, contribute very little to the attainment rating. From (2) we have:

$$0 \leq A_{gs} \leq 1 - \frac{1}{\|P_g\|}$$

The attainment rating $A_{gs} = 0$ if the user has no achievements for game $g$, or in the very unlikely event that everyone else has all the achievements for it. The maximum value $A_{gs} = 1 - \frac{1}{\|P_g\|}$ is achieved in the unlikely event that user $s$ has all achievements for game $g$, and is the only user to have all the achievements. In a traditional rating system, different users may have entirely different standards for what constitutes a given score (e.g. giving out full marks for “I liked this” vs full marks for “this is literally my favorite game of all time”). In contrast, the attainment ratings in (2) are precise, and measure the same degree of attainment for every user. The use of achievements can give scores across the full range for all users; we observe this in Figure 1, with truly high attainment ratings possible, but increasingly rare.

3 GRAPH DATASET

The Steam WebAPI gives access to a wide range of data, including data relating to games, game ownership, friendships, group (e.g. gaming “clans” or social groups representing real-world groups) membership and achievements. Although not explicitly stated, it is reasonable to expect that the data is stored server-side in relational tables rather than in a graph database due to telling limitations in the design of the API. In our work, we propose that using a graph database and querying methods to store and explore the Steam data is appropriate given the contained friendship and group social networks, ownership and achievement data, and size of the overall network. This paper does not detail computational performance or explicit systems implementation of our graph-based recommendation system, with these factors and their deployment in a graph computing system being a current focus of our ongoing research.

The Steam dataset has been analyzed previously, but (to the best of the authors’ knowledge) ours is the first work to use achievement

Figure 1: Distribution of attainment ratings across all users and games in crawled Steam dataset.
data for recommendations and focus on graph structure beyond basic social networks. Some existing papers [6, 7] analyze Steam from a network perspective, but focus on the growth of the friend network, or the propagation of cheaters through interaction. These papers predate the widespread use of achievements data in Steam, and even recent work [27] that provides the crawled dataset online does not include achievement data.

Our crawled dataset is outlined in Table 1. We note that this is merely a subset of the available data that we have chosen to use (and that we were able to crawl before the privacy change of April 10, 2018) for recommendation in this preliminary work.

4 GRAPH-FOCUSED RECOMMENDATION

The use of graph-structured data in recommendation is a well-developed field [4, 13, 19, 30], with numerous techniques designed to incorporate network structural information into recommendation systems. Such work has typically focused on key domain areas such as web search [10, 15, 28] and friend recommendation in social networks [5, 23]. Network models (when used for recommendation) are often fairly basic, with relatively few node and edge types (e.g. only having “users” and “products” as vertex types). The reasoning behind this approach is quite simple; the presence of many different types of nodes and edges can mean that the overabundance of a particular node or edge type can dominate when applying traditional ranking methodologies [3].

In this work, we demonstrate a basic graph query approach for obtaining flexible personalized recommendations using attainment ratings. Instead of focusing on overall rankings, we demonstrate how the richness of the graph structure can be used to obtain recommendations for top quality products that match the users’ specific requirements. We do not cast this as a complete solution to the problem of video game recommendation as such, but rather use it as a proof of concept to demonstrate the value of attainment ratings and querying diverse graph-based datasets.

The social aspect of using (attainment) rating data from actual friends (as opposed to other users designated as “similar” via some clustering or similarity measure) makes particular sense in the context of video games, where online multiplayer is often a big part of a product’s appeal. This suggests the value of the graph-based recommendation in this setting, where users can easily specify queries that account for desired properties and social networks.

We express example product search queries with SQL inspired pseudo-code. For example, Listing 1 is a graph query that says “for a certain user, find the names and purchase costs of five games that are owned by friends, that are developed by companies who have developed games already owned by the user, and order responses according to how the friends ‘rated’ them with respect to attainment”. In this pseudo-code, graph objects are as defined in Table 1. We search in the graph for specific paths of nodes and edges that can succinctly and interpretably describe “complex” queries that actually represent intuitive ideas. Though video games are not in themselves a huge expense in comparison to purchases such as houses, cars etc., we argue that their time investment is large, and that users are thereby willing to spend effort to be precise about their requirements.

Listing 1: Sample graph query pseudo-code

```
SELECT V_G(b). name , V_G(b). cost
PATTERNS V_P(a) − E_F − V_P − E_O − V_G(b)
V_P(a) − E_O − V_G − E_D − V_D − E_D − V_G(b)
WHERE V_P(a). steamid =76561197960653976
ORDERBY AVG(V_P(a)−E_F−V_P−E_O,attainmentRating−V_G(b))
LIMIT 5
```

5 EXPERIMENTAL RESULTS

Here we demonstrate our preliminary experimental results. In this section, we cover two main areas. The first is to summarize the crawl data as it pertains to our derivation of attainment ratings (including the use of attainment ratings in traditional collaborative filtering models). The second is to demonstrate real-world applications of highly customizable graph query recommendations.

5.1 The Value of Attainment Ratings

The plots in Figure 1 show that the distribution of all crawled attainments is very regular, exhibiting similar properties to a Lomax distribution (that is, a Pareto distribution shifted to begin at zero) truncated at the maximum value (see (3)). Indeed, a Kolmogorov-Smirnov test gives a very small KS statistic ($KS = 0.045$) when the ECDF is compared with a $Lomax(shape = 4.78, scale = 0.61)$ distribution. The heavy tail of the empirical distribution can be explained by the existence of “achievement hunters”, those who play in order to attain achievements, rather than the act of playing being an end in and of itself.

Of course, these results relate to the distribution of attainment ratings across all combinations of users and their games. The distribution of attainment ratings varies across developer, genre (see Figure 2) etc. The differences in distribution across these factors provide evidence of commonly held notions within the video game community. For instance, we note that the tails of the Strategy genre are far lighter than the Role-Playing genre, an observation consistent with the idea that excelling in Strategy games requires “talent”, whereas excelling in Role-Playing games is often a simple matter of time investment. We also note that Action games have less mass closer to zero, demonstrating their popularity and tendency for quick satisfaction due to their dense but short nature.

Next, we demonstrate the use of attainment ratings in a collaborative filtering (specifically, SVD++) [21] due to the implicit nature of the attainment ranking data) based recommendation system. Using the default parameters of the Surprise [17] implementation

1 adding (a) and (b) etc. to the patterns identifies that these are the same “objects” in each pattern that must be satisfied.
Figure 2: Density normalized histograms of attainment data grouped by example game genres.

Table 2: RMSE and MAE using SVD++ with 5-fold cross-validation

| Fold | 1     | 2     | 3     | 4     | 5     |
|------|-------|-------|-------|-------|-------|
| RMSE | 0.157 | 0.156 | 0.157 | 0.156 | 0.159 |
| MAE  | 0.116 | 0.116 | 0.117 | 0.115 | 0.116 |

Figure 3: Precision@n and Recall@n for the range of attainment ratings.

(e.g., 20 factors, 20 iterations of SGD with a learning rate of 0.007) of SVD++, we obtained the results shown in Table 2 and Figure 3. It is clear from these results that the attainment ratings well predict other attainment ratings (as shown by Table 2). We also observe from Figure 3 that recommendations made using attainment ratings result in highly relevant recommendations. We argue that the somewhat lower recall values are to be expected from the data; buying many games to be placed into a “backlog” before playing is a common practice.

5.2 The Flexibility of Graph Query Recommendation

The average playtime of a video game is considerably longer than many other forms of entertainment, with games in the Action genre having playtimes typically in the 10-20 hour range, and Role-Playing games often extending into hundreds of hours.

Given the comparatively long length of video games and the large number of products in the dataset, the ability to recall even a small fraction of highly relevant products using attainment ratings suggests an ability to generate a sufficiently large number of recommendations for users for the duration of their use of the platform. This encourages the repeated use of more personalized queries, where a user can narrow these general recommendations to match their specific desires of the moment.

We demonstrate this by means of running Listing 1 for a randomly chosen example user, along with some simple but useful modifications. For instance, we can add the clause “ANTIPATTERNS \( \forall p(a) - \exists vG(b) \) to encode that we do not want to match graph patterns that represent games that the user already owns. These are represented as “Sample Query” and “New Game Query” respectively in Table 3. Note that the attainment ratings as listed are calculated with respect to the immediate neighborhood of the user in the friend network, and not the entire user base. Not only are we explicitly indicating a preference to play games that friends like (and thereby play online with them), trust-based work [11] has shown that recommendations from more distant hops of the social network are less reliable. We observe that four of the five games in Sample Query (all but “Frozen Synapse”) are Action games, with a single Strategy recommendation. When we perform New Game Query, this mix is also noted, with recommendations for both prominent Strategy games (“Magic: The Gathering 2013” and “Total War: Warhammer”) and Action games (“Call of Duty: Black Ops” and “Hitman: Absolution”).

This taste for Strategy games might be noted by the user, and it is a simple matter to restrict results to the Strategy genre (represented by “Strategy Genre Query” in Table 3) by adding the pattern \( \forall vG(b) - \exists vR - \forall vR \) and the WHERE clause condition \( vR.description = \text{Strategy} \).

6 FUTURE WORK

With Steam data now effectively private, we are crawling Xbox Live to serve as a more complete dataset than is possible with our Steam data, allowing us to make increasingly general claims, and improve our measures of performance.

Furthermore, we are developing a graph querying methodology that allows for the efficient specification and implementation of customizable graph queries, such as those that were outlined in Section 5.2. This graph querying methodology, based on probabilistic traversals of the graph, is intended to facilitate a customizable
video game recommendation solution that is capable of scaling to the enormous sizes of the social, purchase, and attainment graphs associated with online videogame marketplaces like Steam and Xbox Live.

This paper has been a practical demonstration of the wealth of marketable information contained within achievement data. We also seek to gain a more theoretical understanding of how to best employ this data for the purpose of recommendation.

REFERENCES

[1] 2002. Xbox Turns Up the Volume on the Future of Video Games. https://web.archive.org/web/20020314071440/http://www.microsoft.com/presspass/press/2002/May02-05-20E3BriefingPR.mspx
[2] 2003. Welcome to Steam. http://store.steampowered.com/
[3] Charu C Aggarwal. 2016. Recommender Systems. Springer International Publishing.
[4] Eugene Agichtein, Carlos Castillo, Debora Donato, Aristides Gionis, and Gildad Mishne. 2008. Finding high-quality content in social media. Proceedings of the 2008 international conference on Web Search and Web Data Mining (2008), 183–194. https://doi.org/10.1145/1341551.1341557
[5] Mohammad Al Hasn and Mohammed J. Zaki. 2011. Linkpredictionin Social Networks. Social Network Data Analysis (2011), 243–275. https://doi.org/10.1007/978-1-4419-8462-3_9 arXiv:arXiv:1011.1669v3
[6] Roi Becker, Yifat Chernihov, Yuval Shavitt, and Noa Zilberman. 2012. An analysis of the Steam community network evolution. In 2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel (IEEEI). 1–5. https://doi.org/10.1109/IEEEI2012.6377133
[7] Jeremy Blackburn, Ramanuja Simha, Nicolas Kourtellis, Xiang Zuo, Clayton Long, Matei Ripeanu, John Skvorcet, and Adriana Iamnitchi. 2011. Cheaters in the Steam Community Gaming Social Network. ArXiv e-prints (2011). arXiv:1112.4915
[8] Chris Bratt. 2018. Why Steam Spy has to close, from the creator himself. (2018). https://www.eurogamer.net/articles/2018-04-11-why-steam-spy-has-to-close-from-the-creator-himself
[9] S. Edsonswan, S. K. Prakasan, D. Kosame, J. Watson, and T. Seymour. 2011. The History of Social Media and its Impact on Business. Management 16, 3 (2011), 79–91. http://search.proquest.com.eprxy.ucd.ie/docview/889149860
[10] Yasuhito Fujwara, Makoto Nakatsuji, Takeshi Yamamuro, Hiroaki Shiokawa, and Makoto Onizuka. 2012. Efficient personalized pagerank with accuracy assurance. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. 15–23. https://doi.org/10.1145/2339530.2339538
[11] Jennifer Ann Golbeck. 2005. Computing and Applying Trust in Web-based Social Networks. Ph.D. Dissertation. University of Maryland. bitstream/handle/1903/2384/umi-umd-2244.pdf?sequence=1
[12] Carlos A. Gomez-Uribe and Neil Hunt. 2015. The Netflix Recommender System. ACM Transactions on Management Information Systems 6, 4 (2015), 1–19. https://doi.org/10.1145/2843948
[13] Marco Gori and Augusto Pucci. 2007. ItemRank: A Random-Walk Based Scoring Algorithm for Recommender Engines. IJCAI 7 (2007), 2766–2771.
[14] Julio Hamari. 2011. Framework for Designing and Evaluating Game Achievements. Proceedings of INGRA 2011 Conference: Think Design Play (2011), 20. https://doi.org/10.1109/224.9966 arXiv:1107.5915
[15] Taher H. Haveliwala. 2002. Topic-sensitive PageRank. In Proceedings of the eleventh international conference on World Wide Web. ACM, 517–526. https://doi.org/10.1145/511446.511513
[16] Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative Filtering for Implicit Feedback. In Proceedings of ICDM. 263–272. https://doi.org/10.1109/ICDM.2008.22 arXiv:arXiv:1208.5721
[17] Nicolas Hug. 2017. Surprise, a Python library for recommender systems. http://surpriselib.com
[18] Mikael Jakobsson. 2011. The achievement machine: Understanding Xbox 360 achievements in gaming practices. Game Studies 11, 1 (2011).
[19] Mohsen Jamali and Martin Ester. 2009. TrustWalker: a random walk model for combining trust-based and item-based recommendation. Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining (2009), 397–406. https://doi.org/10.1145/1551320.1551329
[20] Ben Kuchera. 2017. The anatomy of a review bombing campaign - Polygon. https://www polygons.com/2017/10/4/16418832/psuh-firewatch-steam-review-bomb
[21] Rajeev Kumar, B. K. Verma, and Shyam Sunder Rastogi. 2014. Social Popularity based SVD++ Recommender System. International Journal of Computer Applications 87, 14 (2014), 33–37. https://doi.org/10.5120/15279-4033
[22] Juha Leino and Kari-Jouko Riihia. 2007. Case Amazon: Ratings and Reviews as Part of Recommendations. Proceedings of the 2007 ACM conference on Recommender systems - RecSys ’07 (2007), 137. https://doi.org/10.1145/1297231.1297255