What Makes a Trophy Hunter? An Empirical Analysis of Reddit Discussions

Chien Lu, Jaakko Peltonen, Timo Nummenmaa, Xiaozhou Li and Zheying Zhang
Tampere University

Abstract. In this paper, an empirical data-driven analysis of online discussions of meta-game reward systems is carried out. The data is collected from one of the biggest online discussion forums called Reddit and a text-mining technique called topic modeling is employed. Over 46000 discussion threads from the two most relevant subreddits /r/xboxachievements and /r/Trophies are analyzed and the results of topic modeling shows not only interesting topics but also the (dis)similarity between two text sources the temporal trends of topics. We have found that the volume of related discussions shows an ongoing trend. The topic model results have also revealed that some game genres are more prevalent than others and the (dis)similarities between text sources have been also discovered.

Keywords: trophy hunting, Reddit, topic modeling

1 Introduction

It has been noted that games can also be gamified [10], which has attracted researchers’ attention [9,14]. Early work on applying badge systems to non-game contexts have also used game achievement systems as a source of inspiration [19]. One of the gamification design implementations is a meta-game reward system, which awards visual indicators for completion of tasks and acts as an overarching game in which players obtain rewards through playing across different games [5]. In game industries, meta-game reward systems have been widely used such as badges (STEAM), achievement points (Xbox) or trophies (PlayStation), etc. [12, 15, 30].

Reddit, a major participatory culture platform [17], wh has more than 138,000 active subreddits (each represents a focused sub-community) and more than 330 million users. Due to the number of discussions, the text content has been a valuable resource in various fields [16, 22]. There are two subreddits related to our interests: /r/xboxachievements, a place to “show off achievements” has more than 3 thousand users and /r/Trophies has more than 25 thousand “trophy hunters”. They both yield an abundance of discussions that reflect the gamers or potential gamers’ perception. However, despite containing valuable information, they have not been well explored.

This work performs a large-scale data-driven analysis of gamers’ interests and attitudes toward two reward systems with comprehensive data collection of discussion
threads in the above described two sources. We use topic modeling to computational-
ly extract themes of discussion from the huge corpus. How the prevalence of the ex-
tracted themes varies over the two subreddits and over time are further estimated.

The rest of the paper is structured as follows. Section 2 describes related work.
Section 3 describes data collection and text mining methods. Section 4 describes the
resulting extracted themes. Section 5 provides analysis and discussions. Conclusions,
limitations, and opportunities can be found in Section 6.

2 Related Work

Cruz et al. [5] investigated meta-reward systems for Xbox and Playstation and how
they affect game players’ experience. They collected discussions from focus groups
of 36 console players and used Self-determination theory (SDT), Organismic integra-
tion theory (OIT), Cognitive evaluation theory (CET) and Signaling theory (ST) as
the framework to analyze the discussions. SDT, OIT, CET disclose an individual’s
motivations [26] and ST provides a framework to understand how the information is
sent and interpreted when communicating [6]. Many studies shed light on importance
of game reward systems towards players’ enjoyment and motivation. Jakobsson,
based on study of the Xbox 360 achievement system, indicates it not only works as a
reward system but also as a multiplayer online game in which all Xbox players partic-
ipate compulsorily [11]. Wang and Sun suggest players enjoy game rewards and react
to the motivation they provide while indicating that reward mechanisms foster intrin-
sic motivation and give extrinsic rewards based on relevant psychological theories
[27]. By studying the reward system of World of Warcraft, Rapp emphasizes im-
portance of rewards in shaping player experiences but suggests to consider reward
designs carefully to “surpass the exclusive employment of extrinsic motivators in
gamification design” [24]. Many studies obtained positive outcomes adopting design
similar to game reward systems in gamification and serious games [8, 20].

One example of using empirical data is the work of Wells et al. [28] who analyzed
30,227 PlayStation user profiles to explore how factors including, PlayStation Plus
subscriptions, player regions, etc., influence earning achievements. Using data from
Steam, O’Neill indicates achievements may incentivize more playtime from players
who would have otherwise played less [21]. Players’ achievements data and data min-
ning methods are also used in analyzing their performance in serious games towards
enhancement of players’ motivation and engagement [2]. Many studies have chosen
Reddit as their data source when researching users’ opinions on various topics [13,
29]. We didn’t find works analyzing text of online discussion regarding meta-game
reward systems, perhaps due to difficulties when it comes to a large amount of data.
We tackle the issue by a computer-aided solution explained in Section 3.

We aim to answer three research questions. RQ1. Is discussion of the meta-game
reward system an ongoing trend? RQ2. What kind of games are more mentioned
concerning meta-game reward systems? RQ3. What are similarities & dissimilarities
of discussion between Xbox and PS about the meta-game reward system?
3 Method

Data Collection and Processing. Reddit threads from the two subreddits between 2014 and 2019 (until the 17th of October) were collected via the pushshift API\(^1\). The text was lemmatized and stop words removed. Each submission with its comments was aggregated as a document. We eliminated submissions without comments as they did not evoke discussions. Documents with less than five words were filtered out after processing. In total, we obtained 1384 documents from /r/xboxachievements and 45145 documents from /r/trophies, more details shown in Table 1.

Text Mining. We used a text mining technique called topic modeling to analyze the collected online discussions. It is a probabilistic model that represents each document as a mixture of several latent topics. Each topic speaks for a specific semantic theme within the data collection. The topics are not provided by a human expert but are learned by an algorithm in a data-driven manner. Topic models such as Latent Dirichlet Allocation [3] have been used in domains including game studies (e.g. [7]).

The basic representations found by a topic model are: 1) The strength (prevalence) of the latent topics; each topic has a prevalence in each document, represented as proportions that sum to one, and can be summarized over different data subsets. 2) The vocabulary of the topics, represented as a distribution over typical words in each topic. In this work, a more advanced approach called Structural Topic Model (STM, [25]) is employed. It is a topic model where the strength of topics within a document can depend on document-level covariates, enabling us to model both text content of the documents, and how they are affected by annotations. We take the document source (subreddit) and submission time (initial time of the discussion series) as document-level covariates. The two covariates are taken to model (dis)similarity between two subreddits and the temporal dynamics. Compared to conducting topic modeling in two sources separately (e.g. [1]), this provides an integrated solution via directly treating the effect of the source as a learned parameter in the model.

We use held-out likelihood as the criterion to decide the number of topics $K$, the user-specified parameter of STM. For efficiency we first searched from $K = 10$ to 500 with an interval of 5; the best setting was $K = 355$. We investigated exhaustively from $K = 351$ to 359 and $K = 355$ was still the best. We set up 10 initializing models with $K = 355$; the model with best semantic coherence [18] was adopted as the final model.

4 Results

Table 1 shows the volume of collected data. There are 1384 submissions in /r/xboxachievements and 45145 in /r/trophies. In both subreddits, the number of submissions has grown over time. After topic model training, found topics were labeled under a practice similar to the Thematic analysis [4] but without reading though all the text manually. Labels were decided with authors of this paper examining seman-

\(^1\)https://github.com/pushshift/api
tics of the top words and example documents of the topic (having high prevalence of
the topic). Topics whose top words prominently contained words for specific game
titles or characters were denoted as specific topics about those games; such topics
might concern multiple games, such as a game series, if often discussed together.
While separation of topics on specific games into genres cannot be made fully conclus-
ively, each topic was tagged with one or more genres. Non-game-specific topics
were labeled according to semantic themes they contained. Topics with no clear se-
matic theme were labeled as linguistic topics that provide nuance to the Reddit discus-
sion but which are not a focus of our analysis. Note that each topic may weakly
contain other thematic content besides the main one marked as their label.

Table 1. Number of documents of each year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>xboxachievements</td>
<td>26</td>
<td>154</td>
<td>141</td>
<td>274</td>
<td>294</td>
<td>495</td>
<td>1384</td>
</tr>
<tr>
<td>trophies</td>
<td>2645</td>
<td>4153</td>
<td>6044</td>
<td>7797</td>
<td>11307</td>
<td>13197</td>
<td>45145</td>
</tr>
</tbody>
</table>

The /r/Trophies subreddit introduced a bot ‘PSNTrophyBot’ in 2018 which de-
scribes a game’s trophies, for example, “God of War has 37 trophies: 22 bronze, 8
silver, 6 gold, and platinum...”. The bot adds ‘I am a bot’ to the end of its posts; thus
topics where ‘bot’ was among the top words were considered related to bot activity
and are shown in the online full table but not focused on in later analysis. A selection
of found topics are in Tables 2 - 3. We discuss selected examples next.

Game related topics. 105 of the 355 topics (30% of the topics) turned out to be
game-specific (see online full table). The most discussed game-specific topics were
about games PlayerUnknown’s Battlegrounds, the ‘Infamous’ series, point and click
adventure games from Telltale Games, the ‘Far Cry’ series, Uncharted the Lost Lega-
cy and Conan Exiles, the ‘Assassin’s Creed’ series, the Lego video games series, the
‘God of War’ series for Playstation portable, Demon’s Souls and Bloodborne, and
Rocket League and Teslagrad. We do not present their top words in Tables 2-3 as
they are game-specific: for example, the Telltale Games topic contained words like
cencing games like ‘The Walking Dead’ and ‘Tales from the Borderlands’. However,
examples of game-specific time courses are shown in Fig. 1 to demonstrate the rise
and fall of a game’s popularity for trophy-hunting: PlayerUnknown’s Battlegrounds
(Plot (a), Fig.1) is slowly declining, whereas Grim Fandango and Day of the Tentacle
(Plot (b), Fig.1) have a strong Playstation peak in 2017. See the full set of plots for
other games: e.g. Tomb Raider has little change on Playstation but has a peak around
2016 on Xbox, and Uncharted a broad Playstation peak around 2017. The broader the
peaks, the more enduring the game and its reward mechanism are for trophy hunters;
the strongest ones can serve as examples for future design.

2 Full table at shorturl.at/htH5, all time plots at shorturl.at/ t/uFGN.
<table>
<thead>
<tr>
<th>Topic</th>
<th>Pr(%)</th>
<th>Top 8 Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deciding to try a platinum trophy</td>
<td>1.76</td>
<td>platinum decide tricky huh get trophy twenty gioco</td>
</tr>
<tr>
<td>Time Allocation to Play</td>
<td>1.62</td>
<td>take hour long around say roughly half clock</td>
</tr>
<tr>
<td>Achieving trophies after long time</td>
<td>1.12</td>
<td>get year finally ago almost trophy platinum january</td>
</tr>
<tr>
<td>Easiness and achievability</td>
<td>1.02</td>
<td>easy pretty simple platinum achievable trophy play noby</td>
</tr>
<tr>
<td>Wastefulness and pointless</td>
<td>1.01</td>
<td>like feel matter waste pointless play trophy think</td>
</tr>
<tr>
<td>Games becoming boring or stale</td>
<td>0.95</td>
<td>play everything bore clean throughs normally getstale</td>
</tr>
<tr>
<td>Requesting and appreciating help</td>
<td>0.93</td>
<td>help someone appreciate pleaseable can advance thank</td>
</tr>
<tr>
<td>Congratulations</td>
<td>0.85</td>
<td>congrats thank man congratulation dude jealous madman good</td>
</tr>
<tr>
<td>Earning trophies and happiness</td>
<td>0.69</td>
<td>trophy week earn happy special date favorite talk</td>
</tr>
<tr>
<td>Fun of playing, esp. platformers</td>
<td>0.65</td>
<td>fun blast play platformer platinum definitely cfxtwywmlx battlefield</td>
</tr>
<tr>
<td>Difficulty of getting trophies</td>
<td>0.61</td>
<td>hard say trophy require definitely difficulty platinum take</td>
</tr>
<tr>
<td>Enjoyment and recommendation</td>
<td>0.57</td>
<td>enjoy definitely recommend highly play thoroughly entertain reasonably</td>
</tr>
<tr>
<td>Patches and glitches</td>
<td>0.57</td>
<td>trophy pop problem patch glitched happen glitchy nothing</td>
</tr>
<tr>
<td>Messaging and invites</td>
<td>0.49</td>
<td>send message add request will tomorrow tonight available</td>
</tr>
<tr>
<td>Hunting for trophies</td>
<td>0.38</td>
<td>hunt usually hunter tend generally interest always</td>
</tr>
<tr>
<td>Seek information about a trophy</td>
<td>0.36</td>
<td>else trophy question anyone share hide answer requirement</td>
</tr>
<tr>
<td>Downloadable content and expansions</td>
<td>0.33</td>
<td>dlc base main trophy additional expansion atleast ons</td>
</tr>
<tr>
<td>Progress, corruptions and backups</td>
<td>0.30</td>
<td>save load file progress reload corrupt backup datum</td>
</tr>
<tr>
<td>Dedication required from players</td>
<td>0.28</td>
<td>playerrequire single active dedicate somewhat trophy exclusively</td>
</tr>
<tr>
<td>Relationships</td>
<td>0.26</td>
<td>right wife play girlfriend local trust overcook couch</td>
</tr>
<tr>
<td>Buying in packs for discounts</td>
<td>0.26</td>
<td>buy purchase pack discount platd separately payday consider</td>
</tr>
</tbody>
</table>
Table 3. Top 8 words of selected topics, part 2

<table>
<thead>
<tr>
<th>Topic</th>
<th>Pr(%)</th>
<th>Top 8 Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boss fights and weapon upgrades</td>
<td>0.26</td>
<td>boss fight weapon upgrade defeat fully secret phase</td>
</tr>
<tr>
<td>Shooting and kills</td>
<td>0.26</td>
<td>kill shoot floor shotgun headshots aim rifle</td>
</tr>
<tr>
<td>Fan hopes for sequels</td>
<td>0.25</td>
<td>hope fan sequel disappoint previous quality plot twist</td>
</tr>
<tr>
<td>Purchasing modes and preordering</td>
<td>0.25</td>
<td>version stack digital disc physical cost confuse preorder</td>
</tr>
<tr>
<td>Getting trophies</td>
<td>0.24</td>
<td>another several platted ready chore gratz someday tear</td>
</tr>
<tr>
<td>Milestone</td>
<td>0.22</td>
<td>hit moment milestone onto aswell trophies hot aha</td>
</tr>
<tr>
<td>Frustration and perseverance</td>
<td>0.16</td>
<td>tough frustrate bitch definitely get say perseverance dear</td>
</tr>
<tr>
<td>Enjoyment and motivation</td>
<td>0.16</td>
<td>gaming sometimes always accomplishment enjoyment trophy motivate love</td>
</tr>
<tr>
<td>Multiplayer gaming</td>
<td>0.13</td>
<td>group team pvp randoms chat host teammate invite</td>
</tr>
<tr>
<td>Bugs in games</td>
<td>0.10</td>
<td>bug encounter buggy remix legacy laggy thanatophobia khidr</td>
</tr>
<tr>
<td>Memories and nostalgia</td>
<td>0.09</td>
<td>memory nostalgia hilarious childhood history nostalgic muscle lay</td>
</tr>
<tr>
<td>Attempting and memorizing moves and routes</td>
<td>0.05</td>
<td>try attempt jump second route pause memorize shave</td>
</tr>
</tbody>
</table>

Reward related topics include going for certain types of rewards, earning rewards, reward difficulties, congratulates others, and discussing rewards themselves. Some share feelings of accomplishment, topics Earning trophies and happiness, Milestone, Enjoyment and motivation indicate pride and celebration. In Milestone, a user stated: "...I’d share my 100% of this as I hit level 100 yesterday, the game is consuming my life at the moment!". Some topics imply the achievement difficulty level, e.g. Easiness and achievability and Difficulty of getting trophies; topics like Achieving trophies after long time and Dedication required from players imply perseverance and commitment. In Easiness and achievevability, a user stated: "...I just kept at it until I got em. Otherwise, the trophies in both Pac-Man games are pretty simple...". Topics about feelings (Fun of playing) and fandom (Memories and nostalgia) indicate players’ memorable experiences of the games, and lifestyle (Time allocation to playing, relationship) indicates players’ commitment to playing as part of their life. Wastefulness and pointlessness expresses a feeling of time waste, and Frustration with difficulty and perseverance, Games becoming boring or stale, and Grinding and monotony indicate mixed feelings for the reward system. A user stated in Wastefulness and pointlessness: "...played both games over 200 hours and..."
fully enjoyed them but can never waste time collecting pointless...”. Seek information about a trophy reflects information sharing among users.

Gameplay related topics. These topics indicate players hunting for trophies aim to find information to excel and to be aware of the games’ problems and updates, including gameplay experiences (Multiplayer gaming, Fighting moves: combos, countering and blocking and Boss fights and weapon upgrades) and technical aspects (Patches and glitches, Bugs in games, Saving and restoring progress, corruptions and backups). Enjoyment and recommendation of entertainment and Fan hopes for sequels link individual games to larger entertainment culture. Buying in packs for discounts, Downloadable content and expansions and Purchasing modes and pre-ordering indicate financial planning of trophy hunters.

Temporal Dynamics. Trends of selected topics are in Fig 1. E.g. Grinding and monotony on Playstation peaks in 2016 but is rising on Xbox in 2018-2019; Deciding to try a platinum trophy peaks on Playstation in 2017. Dotted lines represent 95% confidence intervals; non-overlapping intervals of two curves suggest significantly different prevalences. Some topics are Domain-specific. Game-specific topics PlayerUnknown’s BattleGrounds and the “Tomb Raider” series are more prevalent in /r/xboxachievements whereas Grim Fandango, Day of the Tentacle, “Spider-Man” series and Uncharted The Lost Legacy, Conan Exiles are more prevalent in /r/Trophies. Besides, topics Messaging and invites, Completion percentage, Requesting and appreciating help, are found to be more mentioned in /r/xboxachievements whereas Games becoming boring or stale, Gridding and monotony, Time allocation to playing, Deciding to try a platinum trophy and Congratulations are more mentioned in /r/Trophies. Attempting moves and routes has an attention shift, as the trends have opposite directions in the two sources.

5 Discussion

As shown in Table 1, the data volume is drastically growing indicating growth of the communities and impact of meta-game reward systems on gameplay and social media, so we answer our RQ1: the trend is clearly ongoing.

A sizable proportion of topics were game-specific indicating which games are most discussed, answering RQ2. Action-adventure games were the most discussed genre (topics with that genre tag had a total prevalence 0.069), followed by first-person shooters (0.022), role-playing games (0.022) and platformers (0.020). The rest had a prevalence of under 0.010. Discussions on multi-player games (e.g. PlayerUnknown’s Battlegrounds) can be connected to the statement by Jakobsson [11] that the reward system can itself be a massive multiplayer online game. Topics such as Requesting and appreciating help, Congratulations are related to the relatedness in SDT.

The (dis)similarity of focus in the two subreddits is learned by STM and can be evaluated from its parameters, answering RQ3. Topics that are more prevalent in /r/xboxachievements such as Messaging and invites, Completion percentage and Request, appreciate help are more related to pragmatic means of achieving scores. While /r/Trophies has a wider variety of discussions, from perceived emotions
Games becoming boring or stale, Grinding and monotony, and Congratulations to decision making (Time allocation to playing, Deciding whether to try a platinum trophy, and Difficulty of getting trophies). Note that, although the last two topics contain PlayStation-specific words such as “platinum” and “trophy”, other terms (“decide”, “tricky”) can convey general semantic meanings so that they may also appear in /r/xboxachievements threads. This could be due to that PlayStation has different levels of rewards from Bronze to Platinum. This layered design enhances the complexity and potentially enriches the online discussions. The r/xboxachievements subreddit is also smaller by comparison which means some of the discussion on Xbox achievements has spread out more to other venues, such as r/xbox, a more general discussion venue on Xbox related topics.

Fig. 1. Selected topic prevalence over time. Blue: /r/xboxachievements; Red: /r/Trophies; Dot lines: 95% interval.

6 Conclusions, Limitations, and Opportunities

We presented a data-driven topic modeling study of player’s interests and attitudes to reward systems based on two subreddit corpora. Our qualitative analysis of text content was assisted by computational text mining. The results were both interesting on their own and supported human interpretation and efficient analysis of a large amount of text data. We found semantically meaningful topics on game series, trophy-hunting techniques, appreciation of achievements, dissatisfaction with difficulty, grinding, and time use, and their temporal fluctuation. Appreciation or frustration revealed by the topics can help promote a similar appreciation and avoid frustration in future gamification design. The topics provide an understanding of players’ attitudes and opinions
regarding gamification (i.e., badge and achievement systems) of games. The topics show players commit game achievement-related behaviors, e.g., expressing frustration, asking for help, appreciating help, and etc. via online forums. Benefits of such behaviors towards continuous engagement in achievement pursuing and game playing can be useful for other gamification applications.

This work used only Reddit discussions; other large-volume venues include the Playstation Forum and Xbox achievement discussion forum. This can lead to a bias as most Reddit users are from the United States. More categorical analysis can be conducted in future by combining data sources, e.g., Twitter, Steam Community, etc., or comparing to other methods, e.g., Non-Negative Matrix Factorization (NMF) [23].

Deepened understanding of player behavior regarding game trophies and achievements can facilitate the design of personalized gamification, and application of such computational methods towards the understanding of player communities can enrich game studies. We expect the computational approach will inspire similar studies, especially in game studies. The discovered topics and temporal trends can further be used to establish research questions or directions in related fields.

Acknowledgements The work was supported by Academy of Finland decisions 312395 and 313748.

References