

Understanding mobile game success: a study of features related to acquisition, retention and monetization

Átila V. M. Moreira, Vicente V. Filho, Geber L. Ramalho
 Center of Informatics (CIn)
 Federal University of Pernambuco
 Recife-PE, Brazil
 {avmm, vvf, glr}@cin.ufpe.br

Abstract— As mobile game distribution costs gets near zero, the number of available games on app stores, which is already enormous, continues to grow. It gets increasingly difficult for game developers to build a mobile game and achieve the top positions on charts. With it in mind, this paper’s main purpose is to investigate the relationship between game features and the performance achieved by mobile games in terms of number of downloads and gross revenue. A total of 37 game features were analyzed in order to study how each of them influence mobile games’ performance on app stores. The performance of mobile games is measured based on their current position in download and revenue charts on Google Play store. A linear regression model that maps game features and charts performance is trained using a M5 prime classifier and data from 64 mobile games. Results show how each game feature influences the download and revenue performance of successful mobile games.

Keywords—mobile games, app stores, top charts, game design, game features, data mining

I. INTRODUCTION

The shift in gaming industry panorama from premium to freemium business model is recent. It reflects a change in the way people consume and interact with games. Players expect a long-term relationship, and game developers should be ready to provide it. Nonetheless traditional games that are usually distributed through the premium model do not necessarily work and monetize within the freemium model. Therefore, many game developers are facing problems to adapt their team and process to this new baseline.

This paper brings light to the discussion of how to develop games under current panorama. It provides useful insights for game developers and researchers on how game features influence the performance of a mobile game, in both positive and negative ways. We investigate which game features impacts in the success of a mobile game distributed through the Google Play app store.

Literature on this is so far scarce, as it is a new research direction. Most game related analysis try to identify how exactly an aspect of a game relate to the overall game experience; however they do not associate any game aspects

with game performance. The aspects of games that have been considered for analysis and optimization include:

- Research projects that study the effect of narrative on games [1][2][3].
- Studies that examine the relationship between level design parameters of games and player experience [4][5][6].
- Research studies that investigate the development of game rules in a dynamic and automatic way [7][8].
- Articles that investigate the aesthetic side of games and their influence in players’ experience [9][10].

In this paper, we are particularly interested in understanding which features matter to make a successful mobile game. In order to select the right game features that drive monetization, we have chosen a group of features related to the freemium business model. A list of 37 game features was chosen based on previous work on game monetization [16]. These features were used to evaluate the top 100 games in both download and grossing charts of Google Play.

An approach inspired in the CRISP-DM [31] process was used to handle the problem. Firstly, we have acquired and transformed real market data from successful games. The top 100 games on both download and grossing charts were investigated. We have identified the presence, or absence, of 37 game features on each of them. Besides that, we have evaluated their performance - i.e. position on top charts - in the period of a month. There were collected 6,800 entries related to games’ ranking position. After that a regression analysis was conducted to identify the relationship between game features and performance.

The regression method has identified statistically significant correlations. After excluding features with low correlation, we had 6 features that affect download rank and 5 features that affect grossing rank. These features were evaluated by a group of specialists to elucidate and justify results found.

In this work, we present an analysis of which features influence in the success of a mobile game. Our method uses

real data from highly successful mobile games. Firstly we present a list of features related to the freemium business model. These features were used to analyze mobile games on top download and grossing charts of Google Play app store. Their degree of success is measured based on their current position on these charts. A regression analysis was conducted to estimate the relationship between features and their performance. Finally, the results found were appreciated and commented by a group of experts in gaming industry. This paper significantly start a new research area in the gaming industry and present a set of useful insights about what features influence, both in a positive and negative ways, the performance of a mobile game.

II. PROBLEM

Global game industry revenue reached 67 billion dollars in 2012 and is growing at a solid rate of 25% per year. This revenue comes from various sources. There are distinct gaming platforms such as consoles, portable consoles, MMOGs, browser, and mobile devices. However, the latter one is driving a significant part of the industry growth. Nowadays, there are more than 850,000 and 700,000 apps on Apple App Store [11] and Google Play store [12], respectively. One of the main reasons for the success of this platform is the digital distribution of content. It is secure, fast and cheap. It has also influenced the development of new business models.

The most popular business models in game industry used to be the Premium one. It refers to a product that requires users to pay before using. In spite of that, a new business model arose. The freemium is a business model where a game is provided free of charge, but money is charged for advanced features, functionalities or virtual goods. The word "freemium" is a portmanteau neologism combining two aspects of this business model: "free" and "premium".

Nowadays, a total of 69% and 75% of gross revenue from iOS and Android devices, respectively, come from freemium games. Besides that, 98 out of the 100 most profitable games on both platforms - iOS and Android - are freemium games. The mobile game players are willing to spend an average amount of 14 US dollars per transaction (in-app purchase). That is an amount way bigger than the average 0.99-1.99 price tag used on premium games. Among all purchase price points, over 5% of all purchases are for amounts greater than \$50, which rivals the amount paid at retail for top console and PC games.

As distribution costs get nearer to zero, it makes sense to give the game away for free because it allows people to discover them. Giving it for free helps a company get scale. With scale, there are more ways to monetize than a single, up-front transaction. The two main monetization ways are described below:

A. Direct monetization: In-App Purchases (IAP)

This revenue stream is associated with direct sales through purchases of virtual goods inside the game. The most successful games that use IAP as their primary revenue source are listed in the top gross (revenue) list of app stores. The best examples nowadays of great revenue using IAP are the games

Clash of Clans and Hay Day. They generate more than 2.4 million dollars in a daily basis [13].

B. Indirect monetization: Advertising

This form of monetization entails embedding an advertisement into the game. It is intended to attract traffic to a brand, product or service by linking to the advertiser's website. The greater is the number of players, the greater is also the number of ads impressions and clicks. Thus, the revenue from advertising is directly proportional to the user base's size. It is a good estimation to predict that the most successful games that use advertising as a revenue source are listed in the top download list of the app stores. A good example of Advertising revenue is Fruits Ninja of Half Brick. Fruits Ninja generates about 400,000.00 USD per month only in Ads [14].

This change from premium to freemium is recent. In the first semester of 2011, freemium became the dominant business model. It reflected a change in the way people consume and interact with games. Players expect a long-term relationship, and game developers should be ready to provide it. Nonetheless, traditional games that are usually distributed through the premium model do not necessarily work and monetize within the freemium model. Therefore, many game developers are facing problems to adapt their team and process to this new baseline.

This paper brings light to the discussion of how to develop games under current panorama. It provides useful insights for game developers and researchers on how game features influence the performance of a mobile game, in both positive and negative ways.

III. FEATURES

The analysis of a game can involve many different aspects such as aesthetic, gameplay, experience and so on. We are particularly interested in understand which features matter to make a successful mobile game. In order to select the right game features that drive monetization, we have chosen a group of features related to the freemium business model. The ARM funnel (Fig. 1), developed by the research company Kontagent, is commonly used to describe the model. In practical terms, it visualizes mobile gamers passing through a funnel, divided into three stages acquisition, retention and monetization, throughout their lifecycle within the game. With this info in mind we have chosen a group of related features that can be used to indicate how they influence in game performance.

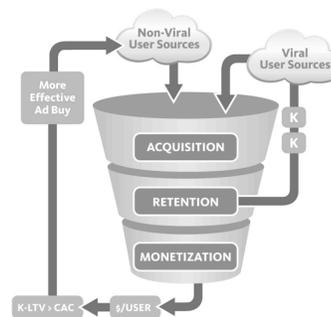


Fig. 1. ARM funnel [15].

The acquisition stage deals with how developers can reach out to users, and acquire players. It deals with features that make game social and viral. The retention stage concerns how to keep players around once they have been acquired. Specifically, it deals with features that make game sticky, or addictive, and is closely related to the gaming mechanics and dynamics presented in gamification. The final stage, monetization, looks at the features used in mobile games to generate revenue from their users. The features should provide incentives for players to pay for the virtual goods.

The original features were extracted from Peter Askelöf's thesis [16], named Monetization of Social Network Games in Japan and the West. They were evaluated and transformed in a more concise list of features. The description of chosen features and executed changes is described below:

A. Social networks

Peter Askelöf [16] demonstrated repeatedly a preoccupation about social features. We can see it on some of his features, such as karma, reputation and friendship points. Usually, social features have great retention power, since they make the game an experience with friends. To facilitate our work divided in 3 major mobile social networks: **Mobage**[ID: R01], **Facebook**[ID: R02] and **Linechat**[ID: R03].

- Mobage - Mobage is a Japanese social network for games, where players can have friends, competition, leaderboard and also connect their mobage account with Facebook.
- Facebook - Facebook is one of the main social networks in the world. It has 1.11 billion unique users per month, or, in other words, more than 15% of all people on earth.
- Linechat - An oriental network with a lot of games and services attached.

B. Social interaction

Between social features there is social interaction. It consists of features like inviting friends or giving friends a gift. We use some of these interactions in this paper: **Request friend help** [ID: AR01] and **Invite friends** [ID: A01].

- Request friend help - this feature forces interaction and usually is used for two main things:
 - To keep players in the game, thus causing retention.
 - To make players call their friends, thus causing acquisition.
- Invite friends - this feature is used to bring other users from social networks to inside the game. Usually it is a feature used by games with social networks.

C. In-App Purchase (IAP)

IAP [ID: M01] is used to buy virtual goods. Usually these virtual goods appear as virtual currency. All games analyzed by Askelöf had IAP feature, due they not differs games with and without this feature. Since we worked with a bigger set of games, we found games with and without this feature.

D. Virtual currency

In freemium games, player commonly uses virtual currency to acquire new content, improvements and virtual goods.

In Askelöf thesis, two types of currency were described: **soft currency** [ID: M02] and **hard currency** [ID: M03]. We decided to use these concepts in this paper.

- Soft currency - this is a type of currency that is earned over time much faster.
- Hard currency – this type of currency has a linear gain during all game time. Usually hard currency is a type of currency that the developer has more control.

E. Gambling

Many games have mini-games with gambling mechanics. For this paper, we needed to know if the game has **gambling** [ID: M04] features, and which type of gambling it is: **hard currency gambling** [ID: M05] or **soft currency gambling** [ID: M06].

F. Game restrictions

A session limit is one way to control the game evolution. A good example of this feature is the energy system of social games. In this paper, we consider all this type of restriction in only one feature. We called this feature **Energy session restriction** [ID: RM01]. This feature helps both retention and monetization.

G. Offers

In the freemium model, a game is no longer a product, but a service. Just like in real life there are commemorative dates with promotions that help in monetization. Askelöf talked about offers involving both content and discounts. In this paper, we simplify this feature thinking only in discounts. However we divide offer in three types: **Unique offer** [ID: M07], **Daily offer** [ID: M08] and **Event offer** [ID: M09].

- Unique offer - usually have a great discount and only can be bought once by the user.
- Daily offer - small discounts that help to convert users to paying users.
- Event offer - This type of offer is improved by commemorative dates like Black Friday and Christmas.

H. Play accelerators

As explained above, there are game restrictions that limit the game session. As presented by Askelöf, some of the things users can buy are consumable **skip times** [ID: M10] and **timed boosts** [ID: M11] to increase session duration. These features's mechanic is very simple:

- Skip time - it is used to make fast build events or deliver content more quickly.
- Timed boost - this feature engages players by a limited amount of time. For example, doubling coins per one day.

I. Upgrades

It is a very common feature in games, since it gives players a sense of progress, which is very important for retention on freemium games. Askelöf combines all types of upgrades on a unique feature that he denominates "Power-ups and boosts". For the purposes of this paper, we divided this feature in three: **Power-up upgrade** [ID: RM02], **Item upgrade** [ID: RM03], **Status upgrade** [ID: RM04].

- Power-up upgrade - upgrades an external element that can somehow improve the score. Example: upgrade the quality of gas that the player gets in level.
- Item upgrade - also an external status, but this time the item can be affected by other items, thus restarting evolution. Example: Improving a sword.
- Status upgrade - upgrade the character's base status. Example: upgrade how high the main character can jump or the amount of life available.

J. Consumable

Askelöf does not mention this on his thesis, but **consumable items** [ID: M12] are a very common feature on games. Consumable items differ from upgrades in the sense that they have a non-persistent effect. Because of that they tend to have a greater impact on monetization, but not so much on retention. A good example of consumable item is a life potion.

K. Customizable

Upgrades and consumables add a great deal of value to games, but on the other hand, **customizable** [ID: M13] does not bring as much value. According to Askelöf, customizable are usually used on social games to allow players to differentiate themselves from others.

L. Level system

It is a very important element responsible for bringing a continuous sense of progress to players. The system is very commonly used on RPGs (Role Playing Games). Two of the features from this game genre that normally appear on mobile games are **Skill trees** [ID: R01] and **Content Unlock** [ID:

R02]. Both of these concepts have the purpose of continuously change the player's gameplay experience.

M. Reward retention

While discussing retention, Askelöf mentions the concept of rewarding retention, which is a way of maintaining a player engaged with the game for a long period of time by rewarding regular players with prizes.

However, there are many ways of rewarding retention, which are: Cumulative **reward retention** [ID: R03], **non-cumulative reward retention** [ID: R04] and **gambling reward retention** [ID: R05].

- Cumulative reward retention - On cumulative reward retention, the player receives an increased prize every day he comes back to the game.
- Non-cumulative reward retention - On non-cumulative, the player always receives the same rewards for coming back to the game regularly.
- Gambling reward retention - On this case, the player receives a random prize by visiting the game on regular time intervals.

N. Punish absence

Askelöf presents two opposing concepts simultaneously, these are rewarding retention and **punishing absence** [ID: R06]. While the former consists of rewarding the user for returning to the game regularly, the latter consists of punishing players who don't return to the game with certain regularity.

O. Gameplay and social interaction

There are four main ways to interact with a game. Askelöf presents only two: competitive and cooperative gameplay, but for this paper we needed to present a more detailed perspective on the existing categories. These categories are: **single play** [ID: R07], **cooperative play** [ID: R08], **competitive play** [ID: R09] and **versus play** [ID: R10].

- Single play - This mode is more common on premium games. The single play usually needs a level based game.
- Cooperative play - Cooperative games are more common on PC, but there are some games that use this mode on mobile, like Clash of Clans or Rage of Bahamut.
- Competitive play - Competitive games, like Subway Surfers, are more common on mobile.
- Versus - Like cooperative, it is more common on Console or PC games, but there are games that implement good versus modes, like Puzzles and Dragons or Song Pop.

P. Achievements and leaderboards

Achievements [ID: R11] are bonus features which are unlocked when players complete certain tasks. They range from simple tasks that players would do anyway, such as completing a level or defeating a boss, to more complicated challenges, such as killing a certain number of enemies in a specific way or completing a level without firing a gun.

Achievements are included within games to extend title's longevity. They also provide players with the impetus to do more than simply completing the game, but finding all of its secrets. Askelöf worked with two types of achievements, but for this work we preferred to simplify this point by only verifying if the game has any type of achievement at all, as we did for leaderboards. **Leaderboards** [ID: R12] are a main factor of competitive games, but simply having a leaderboard does not turn a game into a competitive game, because social interactions are also necessary for that.

Q. Levels

In the same sense that leaderboards are the main factor of a competitive game, **levels** [ID: R13] are the main factor of single player games. However, it is important to note that, even though it is common for single player games to feature levels, this is not always necessary. According to Askelöf: "Levels are an indication of how far a player has progressed in the game. In games such as Ms. Pac-Man, advancement to the next level was clearly visualized in the game by changing they color of the ghosts and the layout of the maze, etc."

R. Random elements

When a game breaks its expected rhythm with a **random element** [ID: R14], it gives players a sense of surprise. Usually, a surprise in a game can break the sense of monotony, so random elements ensure that players will have more interesting experiences, thus increasing retention.

S. Size

Nowadays, the 3G internet is not very responsive. A good game can have problems with user acquisition because of its download size. **Size** [ID: A02] is an important feature, and overall smaller sizes are better. Until the first semester of 2013, Apple blocked any downloads over 3G internet which exceeded the size of 25 MB

IV. METODOLOGY

This section describes the experimental method applied to extract, transform, and analyze data. An approach inspired by the CRISP-DM [31] process was used to tackle the problem. In the first part of methodology, it is described how data was acquired and transformed. After that, a model technique is selected and applied. It is used to identify relationships between game features and performance.

A. Data

In order to evaluate main features that matter to make a successful game, we investigated Google Play's top 100 grossing (revenue) and top 100 downloaded games. AppFigures analytical tool was used to collect data from charts

for the period between April 11 and May 12 of 2013 [30]. The data related to performance (e.g. ranking position) of each game on the top 100 chart was extracted from AppFigures and inserted in a MS Excel spreadsheet. This aimed at facilitating data analysis step, by removing unrelated data.

There were collected 6,800 entries related to games' ranking position. Each entry consists of a row with four columns: game name, download rank position, grossing rank position and date. These entries were associated with ranking positions of 100 games stored between April 11 and May 12 of 2013. Besides ranking position, that shows the games' implicit degree of success; we have also evaluated the games using an extensive questionnaire. The proposed questionnaire consisted of dichotomous questions and used a closed format. Dichotomous questions force respondents to make a choice, e.g. yes/no questions. Closed format is an objective method, excluding the possibility of expressing opinions about games in a free-flowing manner. The questionnaire contained 37 questions related to different aspects of the game. Its purpose was to identify the presence or absence of features.

After business and data understanding phases, data preparation stage started. It covers all activities to construct the final dataset from the initial raw data. Final dataset contains data that will be fed into the modeling tool, e.g. Scores Curves. Activities in this phase include transformation and cleaning of data.

1) Transformation

The 6,800 entries previously collected were grouped by game. Data acquired from the top charts presents 34 entries for each analyzed game. It refers to the specific position on top charts on each day of research time span. A better position in the charts represents an improved degree of success. A game placed in the Top 10 download charts has a much better acquisition model than a game placed between Top 90-100 games. The same analogy is true for grossing (revenue) charts. In 2012, 15% of all revenue on iTunes was generated by games on top 25 grossing list, and the rest of top 100 generated 17% [19]. It is common, however, to use marketing campaigns to improve overall game performance and insert a new game in top charts. It is a useful and smart approach to improve games' visibility and their chance of being chosen by players. The average ranking position was used in order to reduce marketing effects in game analysis and make viable to focus in the game features that improve the game performance.

2) Normalization

The main side effect of former data transformation was the identification of high dispersion rates. It indicates that ranking scores are spread out over a large range of values. It was possible to recognize this effect after sorting games using their average ranking position.

It was necessary to normalize acquired data in a [0,1] range, in order to obtain an easier ranking positions' analysis and comparison. After normalization step, games with zero values are ones with the best performance. Then, these ranking values were inverted so that higher score values indicated better game performance (TABLE 1).

TABLE I. ANALYZED GAMES WITH AVERAGE DOWNLOAD RANK, NORMALIZED AVERAGE DOWNLOAD RANK, AVERAGE GROSSING RANK AND NORMALIZED AVERAGE GROSSING RANK.

| <i>Game Name</i> | <i>Download</i> | <i>Download – Normalized</i> | <i>Grossing</i> | <i>Grossing – Normalized</i> |
|-------------------------------|-----------------|------------------------------|-----------------|------------------------------|
| 3D Bowling | 65,43333 | 0,935437545 | 999 | 0 |
| 4 Pics 1 Word | 5,233333 | 0,995758183 | 35,06667 | 0,96586506 |
| Age of Empire | 77,33 | 0,923517034 | 19,73 | 0,981232465 |
| Angry Birds | 6,83 | 0,994158317 | 99,13 | 0,901673347 |
| Angry Birds Star Wars | 98,93 | 0,901873747 | 13,9 | 0,987074148 |
| Best Casino - Slot Machine | 283,77 | 0,716663327 | 29,83 | 0,971112224 |
| Bike Race Free | 38,1 | 0,962825651 | 60,8 | 0,94008016 |
| Bingo Bash | 51,93 | 0,948967936 | 18,73 | 0,982234469 |
| BINGO Blitz by Buffalo | 83,73 | 0,917104208 | 10,23 | 0,990751503 |
| Bingo Run | 213,0667 | 0,787508317 | 125,3333 | 0,875417535 |
| Blood Brothers (RPG) | 106,43 | 0,894358717 | 4,6 | 0,996392786 |
| Candy Crush Saga | 1 | 1 | 1 | 1 |
| Coin Dozer | 15,97 | 0,985 | 69,6 | 0,931262525 |
| Crime City | 163,17 | 0,83750501 | 18,27 | 0,982695391 |
| Cut the Rope FULL FREE | 29,66667 | 0,971275882 | 999 | 0 |
| D.O.T. Defender of Texel | 126,2 | 0,874549098 | 30,46667 | 0,970474279 |
| Dark Summoner | 239,8 | 0,760721443 | 10,83 | 0,990150301 |
| DH Texas Poker | 14,2 | 0,986773547 | 33,43 | 0,96750501 |
| Diamond Dash | 23,6 | 0,977354709 | 55,67 | 0,945220441 |
| DoubleDown Casino | 62,33 | 0,938547094 | 14,53 | 0,986442886 |
| Drag Racing | 34,5 | 0,966432866 | 51,66667 | 0,949231794 |
| DragonVale | 275,83 | 0,724619238 | 24,97 | 0,975981964 |
| Family Feud | 45,63 | 0,955280561 | 56,17 | 0,944719439 |
| Fantasia | 999 | 0 | 27,27 | 0,973677355 |
| Flow Free | 19,73 | 0,981232465 | 119,37 | 0,881392786 |
| Fruit Ninja Free | 7,27 | 0,993717435 | 999 | 0 |
| Greedy Mouse | 45,63 | 0,955280561 | 999 | 0 |
| HellFire | 186,53 | 0,814098196 | 32,67 | 0,968266533 |
| Hill Climb Racing | 14,83 | 0,986142285 | 51,73 | 0,949168337 |
| Into the Dead | 15,13333 | 0,985838347 | 140 | 0,860721443 |
| Jackpot Slots - Slot Machines | 98,93 | 0,901873747 | 13,9 | 0,987074148 |
| Jetpack Joyride | 40,7 | 0,960220441 | 243,0667 | 0,757448196 |
| Jewels Star | 28,96667 | 0,971977285 | 999 | 0 |
| Jurassic Park Builder | 208,3 | 0,792284569 | 30,93333 | 0,970006683 |
| Kingdoms of Camelot | 999 | 0 | 12,73 | 0,988246493 |
| MARVEL War of Heroes | 999 | 0 | 2 | 0,998997996 |
| Megapolis | 166,97 | 0,833697395 | 12,37 | 0,988607214 |
| Modern War | 196,83 | 0,803777555 | 14,7 | 0,986272545 |

| | | | | |
|-----------------------------|----------|-------------|----------|-------------|
| Paper Toss 2.0 | 49,07 | 0,951833667 | 913,33 | 0,085841683 |
| Pou | 51,4 | 0,949498998 | 229,8333 | 0,770708116 |
| puzzleFree | 34,4 | 0,966533066 | 999 | 0 |
| Rage of Bahamut | 276,83 | 0,723617234 | 4,47 | 0,996523046 |
| Shipwrecked: Lost Island | 117,5 | 0,883266533 | 20,1 | 0,980861723 |
| Shoot Bubble Deluxe | 57,66667 | 0,94321977 | 999 | 0 |
| Slot City - slot machines | 114,5 | 0,886272545 | 10,67 | 0,990310621 |
| Slot Galaxy HD Slot Machine | 151,83 | 0,848867735 | 35,57 | 0,965360721 |
| Slotomania - slot machines | 106,43 | 0,894358717 | 7,33 | 0,993657315 |
| Solitaire | 27,1 | 0,973847695 | 999 | 0 |
| Subway Surfers | 2,4 | 0,998597194 | 22,97 | 0,977985972 |
| Temple Run | 19,83 | 0,981132265 | 999 | 0 |
| Temple Run 2 | 2,93 | 0,998066132 | 86,83 | 0,913997996 |
| TETRIS® free | 50,7 | 0,950200401 | 999 | 0 |
| The Hobbit: Kingdoms | 206,23 | 0,794358717 | 6,83 | 0,994158317 |
| The Simpsons®,ç: Tapped out | 41,93 | 0,958987976 | 4,73 | 0,996262525 |
| The Sims FreePlay | 82,1 | 0,918737475 | 29,03 | 0,971913828 |
| Transformers Legends | 999 | 0 | 27,53 | 0,973416834 |
| Word Search | 39,57 | 0,961352705 | 999 | 0 |
| Words With Friends Free | 19,37 | 0,981593186 | 114,9 | 0,885871743 |
| Zombie Frontier | 34,36 | 0,966573146 | 89,9 | 0,910921844 |
| Zynga Poker | 123,07 | 0,877685371 | 20,47 | 0,980490982 |

3) Cleaning

During normalization step, some games presented missing data in their score position. This occurred because AppFigures analytics tool only stores ranking positions of the first 400 games. Thus, games that do not appear in this Top 400 list on some days were not collected. Only 48 out of 100 analyzed games in top grossing charts presented all daily rankings into the [0,400] range. In top download charts, only 54 out of the 100 analyzed games presented all daily rankings into specified range. All entries with missing data were excluded from the database in order to properly analyze data.

The following histograms present final distribution of ranking position after transformation and cleaning activities. Fig. 2 shows a higher concentration of games on the top download chart's best positions. In Fig. 3, a similar situation occurs, with a higher concentration of games close to top positions, and only a few samples with worse performances.

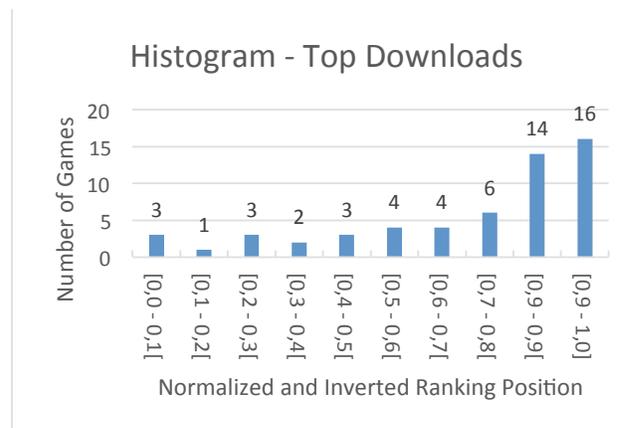


Fig. 2. Histogram of top download games after be normalized and inverted.

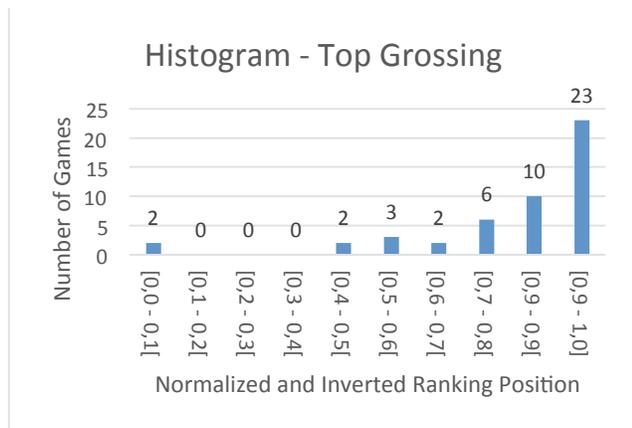


Fig. 3. Histogram of top grossing games after be normalized and inverted.

B. Model

A regression analysis was conducted to help understanding how typical value of dependent variables changes (e.g. position in top charts) when any of independent variables varies, while other independent variables are held fixed. It is a statistical process for estimating relationships among variables. Since dependent variables are continuous, we have chosen a linear regression model.

A linear regression model was built to extract knowledge about relationships among predictive attributes and the class attribute in database. In statistics, linear regression is an approach to modeling the relationship between a scalar dependent variable (e.g. performance of a game) and one or more explanatory variables (e.g. game features). Linear regression calculates regression coefficients that indicate the effect of each explanatory variable over output model. Thus, it makes possible to identify most representatives variables in the analyzed problem.

Data was divided randomly into two groups: training and test groups. Former one is used to produce the model and contains 70% of records. The later one is used to test model's accuracy and contains 30% of the records. During the use of training set to produce the logistic model, the parameters' estimates associated with the variables of the input set are generated. These parameters, also called regression coefficients, are normalized between [-1, +1] and reflect the effect of a given variable in output model.

Linear regression model, as a statistical technique to perform estimation, produces test statistics that can be interpreted using p-values. The p-value indicates the probability that this attribute coefficient is equal to zero (thus proving the null hypothesis). In other words, small p-values correspond to strong evidence. If the p-value is below a predefined limit which is often 0.05, results are designated as "statistically significant".

V. RESULTS

This section describes the main results of experimental methodology applied to real market data. Firstly, early outcomes related to simple statistical data are described.

Average and standard deviation values present useful insights to better understand current industry behavior. Latter, the main results of the linear regression model applied to market data are presented.

As we can see in Table II, both mean and standard deviation are lower in grossing rank. This can be interpreted as a lower mobility in grossing rank. On other hand, download rank shows numbers higher than grossing rank, which means that in download rank games change more quickly.

TABLE II. STATISTIC INFORMATIONS

| Type | Average | Standard Deviation |
|----------|---------|--------------------|
| Download | 32.01 | 38.52 |
| Grossing | 18.50 | 10.60 |

Linear regression's results are presented in Tables III and IV and Fig. 4 and 5. They show the variables list and their parameters. Attributes in tables have been sorted by p-value, with lower values indicating greater significance for this variable. Besides that, we have omitted non-statistically significant results for the sake of clarity. Coefficient values represent how each variable influence the performance of a game in both download and grossing charts.

The coefficient absolute value, or modulus, refers to the intensity of influence on performance outcomes. A greater value means that the feature presents a higher leverage. Furthermore, coefficient value sign - i.e. positive or negative - is related to how each feature influence on game performance. A positive value means that the analyzed feature influence overall game performance in a positive way. The opposite is also true. A negative value means that the investigated feature affects game performance in a negative way.

The tables present also an intercept attribute. It is a constant representing line intercept with vertical axis. In other words, it indicates that even when all of attributes are zero, there will be some amount of positive influence. It makes sense since all analyzed games are in top 10 charts and present a good performance.

TABLE III. GROSSING RESULTS

| Attribute | Coefficient | Std. Error | t-Stat | p-Value |
|------------------------|-------------|------------|--------|---------|
| (Intercept) | 0.5372 | 0.0982 | 5.467 | ~0.0000 |
| Customizable | -0.4134 | 0.0817 | -5.055 | ~0.0000 |
| Random elements | 0.2197 | 0.0658 | 3.336 | 0.0036 |
| Event offer | 0.2206 | 0.0745 | 2.959 | 0.0086 |
| Versus | 0.1930 | 0.0803 | 2.403 | 0.0296 |
| Soft currency gambling | -0.1798 | 0.0788 | -2.281 | 0.0386 |

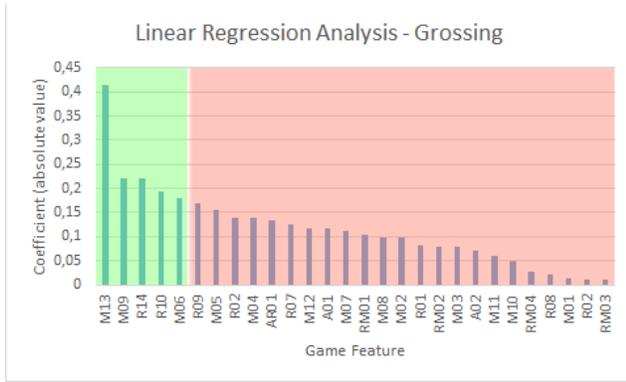


Fig. 4. Histogram of top grossing games using coefficient. Game features are represented by ID presented on session III.

| Attribute | Coefficient | Std. Error | t-Stat | p-Value |
|-------------|-------------|------------|--------|---------|
| upgrade | | | | |
| Mobage | 0.2091 | 0.0926 | 2.257 | 0.0359 |
| Achievement | 0.1039 | 0.0483 | 2.152 | 0.0455 |

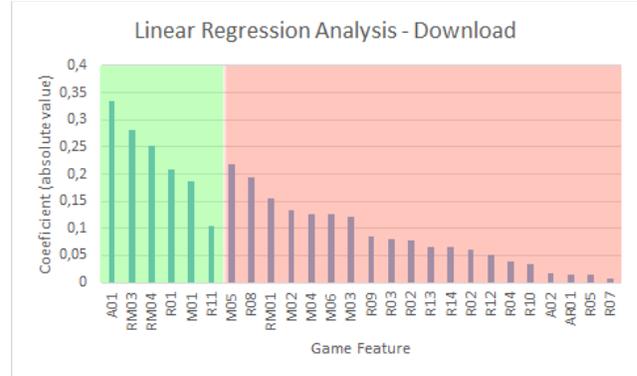


Fig. 5. Histogram of top download games using coefficient. Game features are represented by ID presented on session III.

TABLE IV. DOWNLOAD RESULTS

| Attribute | Coefficient | Std. Error | t-Stat | p-Value |
|------------------------|-------------|------------|--------|---------|
| (Intercept) | 0.8436 | 0.0684 | 12.328 | ~0.0000 |
| Invite friends feature | -0.3343 | 0.0875 | 4.35 | ~0.0000 |
| IAP | 0.1857 | 0.0427 | -3.819 | ~0.0000 |
| Item upgrade | -0.2807 | 0.0787 | -3.564 | 0.0013 |
| Status | -0.2519 | 0.1012 | -2.489 | 0.0208 |

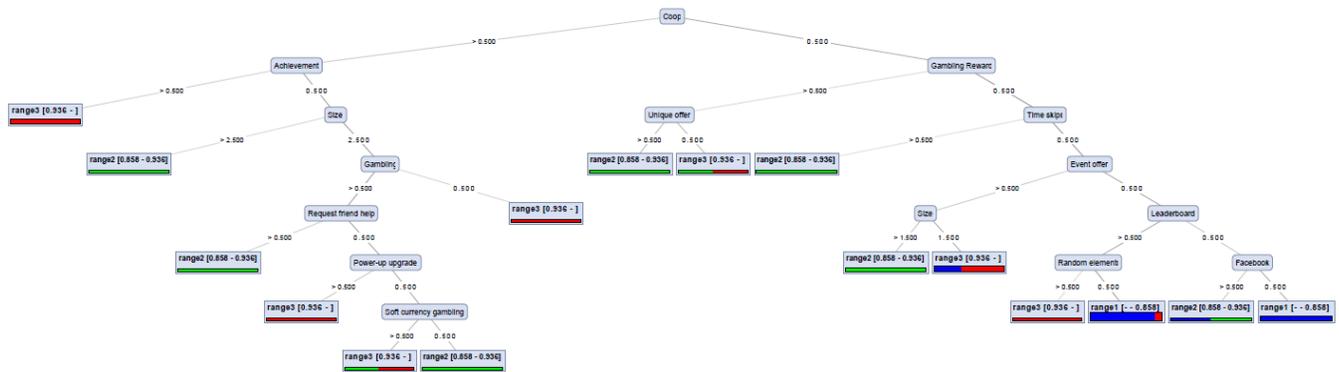


Fig. 6. Decision tree generated by Rapid Miner to Grossing Rank.

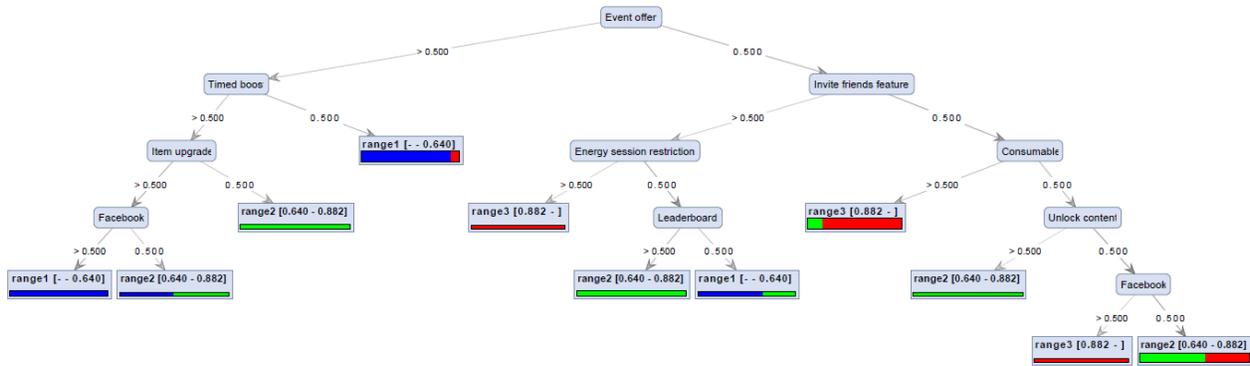


Fig. 7. Decision tree generated by Rapid Miner to Download Rank.

VI. ANALYSIS

This section discusses main findings of statistical approach used to analyze which features influence mobile games' performance. Statistically significant correlations were evaluated by a group of specialists to elucidate and justify results found. Producers and designers with more than 10 years of experience in game development form the team of experts.

After excluding features with low relationship discovered by p-value, we had 6 features that affect download rank and 5 features that affect grossing rank. The table V shows these features and their type of effect. A positive sign means that the feature is related to a better game performance. On the other hand, a negative signal means exactly the opposite.

TABLE V. ALL RESULTS

| Download | | Grossing | |
|-------------|----------------|----------------|------------------------|
| + | - | + | - |
| Achievement | Invite friends | Rand. Elements | Customizable |
| Mobage | Item upgrade | Event offer | Soft currency gambling |
| IAP | Status upgrade | versus | |

With the help of game specialists, we discuss each of identified representative features that affect games performance. Initially, features associated with games in the top download charts were evaluated. After that, features related with games in the top grossing charts were examined.

A. Download

Download ranks is usually affected by external marketing campaigns. However, there are cases when only a small portion of gamers comes from these sources. In the Pearl's Peril game, developed by Wooga, only 10% of its user's base is acquired from marketing campaigns. Remaining 90% comes from social resources such as virality and cross-promotion [20]. With this information in mind, we infer relationships to clarify how each feature affects download rank.

1) Achievement

This feature influences games in a positive way. The achievement feature has presented a coefficient of +0.10396. In terms of game experience, achievements are a common way to recognize players for their performance. Besides that, players usually win not only a reward (e.g. a rare virtual good) but also a virtual trophy, medal or badge. In this specific moment, players are more than willing to share their awards as feed in social networks - this way increasing the game's visibility. Aaron Greenberg, group product manager for Xbox 360 and Xbox Live, said in an interview with The Hollywood Reporter: "We see gamers coming back to us because we give points, other platforms don't" [21].

2) Mobage

This feature influences games in a positive way - Mobage feature presented a coefficient of +0.2091. Games with Mobage use cross promotion in their giant user base, counting with more than 49 million players [22]. Also, these games force users to interact with others users. Thus, it makes sense for this feature to help on download numbers. Boney The Runner game, for instance, was launched on October 2012 with this feature and has achieved a half million players in less than one year [23]. This game has spent no money in their marketing efforts.

3) IAP

This feature influences games in a positive way. IAP feature has presented a coefficient of +0.1857. Mobile games that present IAP usually offer a lot of content because it stimulates virtual goods consumption. Games that have IAP usually have more content insofar as player pay to consume more content. According to Michail Katkoff [24], product manager at Supercell, IAP is an option to non-virality. In other words, players that do not want to use IAP need to realize one of two following options: grind or virality. The virality action, mentioned above, helps the acquisition, in other words download rank.

4) Invite Friends

This feature influences games in a negative way - it has presented a coefficient of -0.3343. This result was unexpected since it is a typical viral feature.

5) Item Upgrade

This feature influences games in a negative way, presenting a coefficient of -0.2807. We expected this feature to help both download and grossing, since this feature implies in more content. Unfortunately this works shows an opposite result, and we will investigate more about this in future opportunities.

6) Status Upgrade

This feature influences games in a negative way - it has presented a coefficient of -0.2519. Like Item Upgrade, we also expected this feature to increase both ranks. We will also investigate more about this in the future.

B. Grossing

Grossing ranks is only affected by quantity of purchases inside the application. Now, we will discuss about some features that affect this, however to do so we need to understand how the receipt is changed.

We can admit Average Receipt Per User (ARPU) is a function of receipt by number of users (1) and we can also admit that average receipt per user is a function of Conversion Rate(CR) and Average Receipt Per Paying User (ARPPU) (2). Then, to constant downloads number, we have two aspects that affects receipt (3): ARPPU and CR [25].

$$ARPU = \frac{Receipt}{Downloads} \quad (1)$$

$$ARPU = ARPPU * CR \quad (2)$$

$$Receipt = ARPPU * CR * Downloads \quad (3)$$

1) Random elements

Based on our statistical analysis, this feature influences games in a positive way - it has presented a coefficient of +0.2197. Random elements are usually a signal that the game has good content. A game with great quantity of content implies more things to spend real money in, this way increasing ARPPU [26].

2) Event offer

This feature influences games in a positive way - it has presented a coefficient of 0.2206. Event offers are temporary reduction of prices or exclusive content for a short amount of time. Players that are near to convert from free to paying players, usually convert in these moments, thus increasing CR [27]. Due to this, this feature helps monetization. In our games, we noticed a drastic change on income during commemorative dates, especially when we make daily offers.

3) Versus

This feature influences games in a positive way - it has presented a coefficient of 0.1930. Versus is a strong feature. Depending on how it is done, it can instigate competition and improve CR. According to Michail Katkoff [25], comparing score and status with friends can help the CR.

4) Customization

This feature influences games in a negative way - it has presented a coefficient of -0.4134. It makes sense since pure customization do not bring value to core gameplay. According to Valadares [28], only 2% of freemium revenue comes with customizations. Flurry gets data from more than 2 million players every day.

5) Soft currency gambling

This feature influences games in a negative way - it has presented a coefficient of -0.1798. This occurs due to the fact that in this system the player can get some premium content without pay, this way decreasing CR. We tested this insight in one of our games, Boney The Runner, that used soft currency gambling. We changed for another type of gambling and game revenue really increased.

VII. CONCLUSIONS AND FUTURE WORK

This paper presented an evaluation of which features matters to build a successful game. Data analysis demonstrated a relationship between some features and the performance of their games on both download and grossing (revenue) charts on Google Play. A linear regression approach has exposed the degree of relationship and also the confidence interval of each game feature. The main findings are how some feature affect the game performance on the store. Regarding download rank, we found that features as Mobage, IAP and Achievement affect the game positively. On the other hand, features like Invite friend, Upgrade item and Upgrade status affect negatively. In the case of Grossing rank, we found that features like Random elements, Event offer and Versus affects the grossing rank positively, whereas Customization and Soft currency gambling affect negatively.

In our future work, we will acquire more data to expand results of the presented work. We are going to extract data also from games with poor performance. With this information, it will be possible to not only understand the relationship between features and game performance, but also develop a binary classification model. In this case, the classification model could be used to identify the successful chance for a given game. This approach can also be used to suggest a feature for a game based on the associated increase in their performance probability.

ACKNOWLEDGMENT

The authors would like to thank our sponsors Big Hut Games [29], Manifesto Games Studio [32] and Capes [33].

REFERENCES

- [1] Roberts, David L. "Computational techniques for reasoning about and shaping player experiences in interactive narratives", Georgia Institute of Technology, April 6, 2010.
- [2] H. Jenkins. "Game Design as Narrative Architecture", anabiosispress.org. [Online]. Available: http://www.anabiosispress.org/VM606/1stPerson_hjenkins.pdf [Accessed: July 12, 2013].
- [3] R. Michael Young and Mark Riedl. "Towards an architecture for intelligent control of narrative in interactive virtual worlds." In Proceedings of the 8th international conference on Intelligent user interfaces (IUI '03). 2003.

- [4] Chris Pedersen, Julian Togelius and Georgios N. Yannakakis. "Modeling Player Experience in Super Mario Bros". 2009.
- [5] Gillian Smith, Mike Treanor, Jim Whitehead, Michael Mateas. "Rhythm-Based Level Generation for 2D Platformers". 2009.
- [6] Peter Mawhorter, Michael Mateas. "Procedural Level Generation Using Occupancy-Regulated Extension".
- [7] J. Togelius, Lugano IDSIA, J. Schmidhuber. "An experiment in automatic game design". 2008.
- [8] Mark J. Nelson, Michael Mateas. "Towards Automatic Game Design". 10th Congress of the Italian Association for Artificial Intelligence, Rome, Italy, pp. 626-637, October 2007.
- [9] D. Myers. "A Q-Study of Game Player Aesthetics". *Simulation & Gaming*, pp. 375-396, December 1990.
- [10] Thomas W. Malone. "Heuristics for designing enjoyable user interfaces: Lessons from computer games." In *Proceedings of the 1982 Conference on Human Factors in Computing Systems (CHI '82)*. ACM, New York, NY, USA, 63-68. 1982.
- [11] M. Rodrigues. "App Store da Apple atinge marco histórico de 50 bilhões de downloads", *Apple.com*, May 16, 2013. [Online]. Available: <http://www.apple.com/br/pr/library/2013/05/16Apples-App-Store-Marks-Historic-50-Billionth-Download.html> [Accessed: July 12, 2013].
- [12] B. Womack. "Google Says 700,000 Applications Available for Android", *businessweek.com*, October 29, 2012. [Online]. Available: <http://www.businessweek.com/news/2012-10-29/google-says-700-000-applications-available-for-android-devices> [Accessed: July 12, 2013].
- [13] K. Strauss. "The \$2.4 Million-Per-Day Company: Supercell", *forbes.com*, April 18, 2013. [Online]. Available: <http://www.forbes.com/sites/karstenstrauss/2013/04/18/the-2-4-million-per-day-company-supercell/> [Accessed: July 12, 2013].
- [14] J. D. Rey. "Mobile App Economics: 'Fruit Ninja' Makes \$400,000 a Month on Ads", *Adage.com*, July 11, 2013. [Online]. Available: <http://adage.com/article/digital/mobile-app-economics-fruit-ninja-makes-400-000-a-month-ads/235965/> [Accessed: July 12, 2013].
- [15] Kontagent, "The top 7 metrics of high Successful Social Companies", *kontagent.com*. [Online]. Available: http://static.kontagent.com/whitepaper/KNT_WP_Top7Metrics_P3_fina_lx.pdf [Accessed: July 12, 2013].
- [16] P. Askelöf, "Monetization of Social Network Games in Japan and the West", Lund University, Faculty of Engineering, LTH, 2013.
- [17] Facebook Inc., "Number of active users at Facebook over the years". [Online]. Available: <http://news.yahoo.com/number-active-users-facebook-over-230449748.html> [Accessed: July 12, 2013].
- [18] M. Katkoff, "Primed to spend", *deconstructoroffun.blogspot.com*, May 13, 2012. [Online]. Available: <http://deconstructoroffun.blogspot.com.br/2012/05/primed-to-spend-slowingdown-rate-of.html> [Accessed: July 12, 2013].
- [19] M. E. Gordon, "The History of App Pricing. And Why Most Apps Are Free", *flurry.com*, July 18, 2013. [Online]. Available: <http://blog.flurry.com/?Tag=App%20Revenue> [Accessed: July 12, 2013].
- [20] E. Maiberg, "Pearl's Peril is Wooga's fastest-growing game to date", *insidesocialgames.com*, May 13, 2013. [Online]. Available: <http://www.insidesocialgames.com/2013/05/13/pearls-peril-is-woogas-fastest-growing-game-to-date/> [Accessed: July 12, 2013].
- [21] M. Jakobsson, "The Achievement Machine: Understanding Xbox 360 Achievements in Gaming Practices", *gamestudies.org*. [Online]. Available: <http://gamestudies.org/1101/articles/jakobsson> [Accessed: July 15, 2013].
- [22] DeNA. [Online]. Available: <http://developer.dena.jp/mbga/en/> [Accessed: July 15, 2013].
- [23] Mobage. [Online]. Available: <http://mobage.com/games/boney-the-runner> [Accessed: July 15, 2013].
- [24] M. Katkoff, "There's More Than Just the Money...", *deconstructoroffun.blogspot.com*, May 5, 2012. [Online]. Available: <http://deconstructoroffun.blogspot.com.br/2012/05/theres-more-than-just-money.html> [Accessed: July 12, 2013].
- [25] M. Katkoff, "From Players to Payers", *deconstructoroffun.blogspot.com*, May 4, 2012. [Online]. Available: <http://deconstructoroffun.blogspot.com.br/2012/05/from-players-to-payers-4-steps-to.html> [Accessed: July 12, 2013].
- [26] M. Handrahan, "Content is king: the growing pains of free-to-play", *gamesindustry.biz*, November 8, 2012. [Online]. Available: <http://www.gamesindustry.biz/articles/2012-11-08-content-is-king-the-growing-pains-of-free-to-play> [Accessed: July 12, 2013].
- [27] M. Katkoff, "Three Steps from Paid to Freemium", *deconstructoroffun.blogspot.com*, June 9, 2012. [Online]. Available: <http://deconstructoroffun.blogspot.com.br/2012/06/three-steps-from-paid-to-freemium.html> [Accessed: July 12, 2013].
- [28] J. Valadares, "Freemium Mobile Gamers Spend Most Money on Items They Don't Keep", August 16, 2011. [Online]. Available: <http://blog.flurry.com/bid/70096/Freemium-Mobile-Gamers-Spend-Most-Money-on-Items-They-Don-t-Keep> [Accessed: July 12, 2013].
- [29] Big Hut Games. [Online]. Available: <http://www.bighutgames.com>
- [30] Appfigures. [Online]. Available: <http://www.appfigures.com/>
- [31] The CRISP-DM process model. [Online]. Available: <http://www.crisp-dm.org/>. 1999. [Accessed: April 9, 2013]
- [32] Manifesto Games Studio. [Online]. Available: <http://www.manifestogames.com.br>
- [33] CAPES. [Online] Available: <http://www.capes.gov.br/>