Game Bot Detection in Online Role Player Game through Behavioural Features

Mario Luca Bernardi[1], Marta Cimitile[2], Fabio Martinelli[3], Francesco Mercaldo[3]

[1]*Giustino Fortunato University, Benevento, Italy*

[2]*Unitelma Sapienza University, Rome, Italy*

[3]*Institute for Informatics and Telematics, National Research Council of Italy (CNR), Pisa, Italy*

m.bernardi@unifortunato.eu, marta.cimitile@unitelma.it, {fabio.martinelli, francesco.mercaldo}@iit.cnr.it

Keywords: game bot, machine learning, game bot detection, security, testing

Abstract: The market of online games is highly increasing in the last years and thanks to the availability of always more effective gaming infrastructures and the increased quality of the developed games. The diffusion of on line games also increases the use of game bots to automatically perform malicious tasks obtaining some rewards (with a consequent economical advantage or popularity) in the game community with low effort. These causes the disappointment of the game players community becoming a critical issue for the game developers. For this reason, distinguishing between game bots and human behaviour being became essential in order to detect the malicious tasks and consequently increase the players satisfaction. In this paper authors propose an approach to the game bot detection in the online role player games based on the adoption of machine learning techniques in order to discriminate between users and game bots basing on some user behavioral features. The approach is applied to a real-world dataset of a popular role player game and the obtained results are encouraging.

1 INTRODUCTION

With the increasing growing of network connections working at high-speed, several activities are every-day performed using online services. As a matter of fact, more and more people use Internet network to perform a plethora of operations like buying the train tickets, making a bank check or taking a meeting reservation with a public/private office (Bernardi et al., 2012). This trend is also extended to the worldwide games market that in last years is enriched by the development of several platforms for online gaming and by the increasing quality of the available games. An online game is a video game that is either partially or primarily played through the Internet network or another computer network (Adams, 2014). Online games are ubiquitous on modern gaming platforms, including PCs, consoles and mobile devices, and span many genres, including first-person shooters, strategy games and massively multiplayer online role-playing games (MMORPG) (Quandt and Kröger, 2013). The design of online games can range from simple text-based environments to the incorporation of complex graphics and virtual worlds (Seay et al., 2004). The prominence of online components within a game can range from being minor features, such as an online leaderboard, to being part of core gameplay, such as directly playing against other players. Many online games create their own online communities, while other games, especially social ones, integrate the existing real-life communities of real players (Wellman and Gulia, 1999). Online games have attracted players from a variety of ages, nationalities, and occupations (Yee, 2008; Griffiths et al., 2004). This represents an appealing scenario for attacker in order to perpetrate illegal activities in the online world (Chen et al., 2004). Most illegal activities occur continuously because cyber assets, such as game items and cyber money in online games, can be changed into real currency (Paulson and Weber, 2006). Moreover, attackers can be also interested to capture reserved information about the human users that access to the game resources or to became popular in the game community. The malicious activities are generally performed using some game bots in order to repeat continuously a malicious task in the online game. A game bot is an artificial intelligent system software that plays a video game in the place of a human (Yampolskiy and Govindaraju, 2008). Game bots are used in a variety of video game genres for a variety of tasks: a bot written for a first-person shooter (FPS) works very differently from one written for a

MMORPG. The former may include analysis of the map and even basic strategy; the latter may be used to automate a repetitive and tedious task. Basically a game bot is an automated program that plays a given game on behalf of a human player. Game bots can earn much more game money and items than human users because the former can play without requiring a break. Moreover game bots are also used in order to disturb human users because they consistently consume game resources (for instance, game bots defeat all monsters quite rapidly and harvest items, such as farm produce, before human users have an opportunity to harvest them). Accordingly, game bots cause complaints from human users and damage the reputation of the online game service provider. Furthermore, game bots can cause inflation in a game's economy and shorten the game's lifecycle, which defeats the purpose for which game companies develop such games (Kang et al., 2016).

Starting from these considerations, in this paper we propose a method able to detect game bots in a MMORPG environment. The proposed method is based on the adoption of machine learning techniques to build several classifiers able to discriminate between users and game bots basing on behavioral features: Player Information, Player Actions, Group Activities, Social Interaction Diversity and Network Measures. The paper poses the following research question:

> Is it possible to detect game botnet using the proposed behavioural features with the aim of evaluating the effectiveness of discriminating human and botnet behaviours from the viewpoint of the investigators in the context of a controlled experiment on a real MMORPG game?

We evaluate the effectiveness of our method using a real-world dataset obtained from the operation of Aion: The Tower of Eternity [1] game, a MMORPG fantasy free-to-play popular game. The dataset contains operations performed by real players and game bots, identified and labeled by the game company. The rest of the paper is organized as follows: the next section provides an overview of related work; the following section illustrates the proposed features and the detection technique; the fourth section presents the results of the evaluation, and, finally, conclusion and future works are given in the last section.

[1] https://en.aion.gameforge.com/website/

2 RELATED WORK

In the last years, an high number of studies have been pointed towards the adoption of machine learning techniques to identify game bots in the MMORPG. In (Kim et al., 2005), an approach to the bot detection based on the analysis of the window event sequences produced by the game players, is proposed. Here learning algorithms are used to classify and distinguish the event sequences (transformed as a set of attributes) coming out from human or auto player. Authors in (Chen et al., 2008) propose a manifold learning approach for game bots detection. The approach analyzes the avatars movement trajectories basing on the assumption that there is a difference between human players trajectories and those of game bots. In (Chung et al., 2015), an approach consisting to compare the traffic generated by human players versus the game bots is proposed. While the discussed approaches are mainly based on global model to discriminate between human and bot net, some more recent approaches are focused on the behavioral analysis of the game playing (Thawonmas et al., 2008; Kashifuji, 2008; Hilaire et al., 2010; Mishima et al., 2013; Kang et al., 2016) considering the different game style of the human player as a discriminant factor with respect to the game bots. For example, (Oh et al., 2013) assumes that humans and game bots tend to form their social network in different ways. According to this some features are used to capture this social behavior for detecting game bots. The discussed behavioral approaches drawback consists in the limited number of observed behavioral features (one or two behavioral features) that are more related to the game domain than to the player game style. This reduced number of features make the approach dependent from the analyzed game. This issue is overrun in (Chung et al., 2015) where authors propose an approach based on the behavior analysis of the game playing. This approach is very similar to our proposed approach basing on both game-related and player-related features. With respect to (Chung et al., 2015), in our proposed approach the obtained results show a higher precision of our proposed approach in the bot net detection.

In addition in order to demonstrate that the developed method is useful to detect game bot in the real environment we perform a feature selection step, in order to consider in the classification phase only the most discriminative features between human users and game bot, and we discuss the time employed to build the classifiers representing the maximum temporal interval to detect a game bot behaviour by our method.

3 THE METHOD

This section the approach we proposed is discussed in detail.

As we stated into the introduction, the aim of the following study is to demonstrate that the behaviour of a game bot is different from the human user one. To this end, we analyse a set of behavioural features from both human users and game bots. These set of features values were extracted and computed by a real game company as described as described in (Kang et al., 2016). We consider some features related to the players and other features related to the game.

In particular we grouped the behavioural features in following categories: Player Information (PI), Player Actions (PA) (they are player related features) and Group Activities (GA), Social Interaction Diversity (SID) and Network Measures (NM) (they are game related features).

Table 1 shows the considered features for each category we considered.

Considering that the aim of game bots is to automatically perform some game tasks, we observe PI features in order to find a gap between the values of the personal features of game bots and those of human users. According to this we consider the following features: login frequency (PI_1), play time (PI_2), game money (PI_3) and number of IP address ((PI_4) . Similarly, we evaluate the PA features containing some features like sitting (PA_1) (i.e., an action taken by players to recover their health), earning experience points (PA_2), obtaining items (PA_3), earning game money (PA_4), earning player kill points (PA_5), harvesting items (PA_6), resurrecting (PA_7), restoring experience points (PA_8), being killed by a non-player and/or player character (PA_9), and using portals (PA_{10}). We investigate whether these actions can reflect the behavioral characteristics of game bots and human users. For instance, game bots sit more frequently than human users in order to recover health and many points. Moreover, a player can acquire player kill points by defeating players of opposing factions. Player kill points can be used to purchase various items from vendors. Player kill points are also used to determine a players rank within the game world. In the Aion game[2], the more player kill points a player has, the higher is the rank of the player. The high ranking player can feel a sense of accomplishment. In addition game bots often connect to the game and can play for 24 consecutive hours, differently from human users, that typically are not able to play during several time-windows, for instance during work and sleep ones. Considering that usually when a user reaches a certain rank level he obtains powers to fight the enemies more effectively, game bots can more easily obtain these powers, and the attacker once their characters have obtained the powers resell them to other users [3].

The rationale behind the GA feature is that there is a gap between the values of the social features of game bots and those of human users because game bots do not attempt to social as humans. As matter of fact, as a consequence of the game bot protract play time, we expected that the GA category is different between game bots and human users.

The GA category includes the average duration of party play (GA_1) and number of guild activities (GA_2). Party play is a group play formed by two or more players in order to undertake quests or missions together. The goals of party play commonly are to complete difficult quests by collaboration and enjoy socialization. Interestingly, some game bots perform party play, but the goal of party play of the game bots is different from that of human users. Their aim is to acquire game money and items faster and more efficiently in order to obtain powers. For these reason, we expected that there are the behavioral differences between game bots and human users.

The SID features indicate the entropy of party play. Game bots concentrate only on particular actions, whereas human users execute multiple tasks as needed to thrive in the online game world.

Relating to the NM category, the player's social interaction network can be represented as a graph with characters as the nodes and interactions between them as the edges. An edge between two nodes (players) in this graph may, for example, highlight the transfer of an item between the two nodes. The features of NM include the degree centrality (NM_1), betweenness centrality (NM_2), closeness (NM_3), eigenvector centrality (NM_4), eccentricity (NM_5), authority (NM_6), hub (NM_7), PageRank (NM_8), and clustering coefficient (NM_9).

The NM category features definition is explained in Table 2.

4 THE EVALUATION

We designed an experiment in order to evaluate the effectiveness of the feature vector we propose, expressed through the research question RQ stated in the introduction. More specifically, the experiment is

[2] https://sites.google.com/a/hksecurity.net/ocslab/Datasets/game-bot-detection

[3] http://www.geek.com/games/how-one-diablo-3-player-pulled-in-130000-from-the-real-money-auction-house-1601959/

Category	Features
Player Information	PI_1, PI_2, PI_3, PI_4
Player Actions	PA_1, PA_2, PA_3, PA_4,PA_5, PA_6, PA_7, PA_8, PA_9, PA_{10}
Group Activities	GA_1, GA_2
Social Interaction diversity	SID_1
Network measures	NM_1, NM_2, NM_3, NM_4, NM_5, NM_6, NM_7, NM_8, NM_9

Table 1: The features involved in the study with the correspondent category.

NM category feature	Description
Degree centrality	This features represents the centrality focused on the degree. The more edges an actor has, the more important it is
Betweenness centrality	It counts the number of shortest paths between two nodes on which a given actor resides
Closeness centrality	An actor is considered important if it is relatively close to all other actors. Closeness is based on the inverse of the distance of each actor to every other actor in the network
Eigenvector centrality	Indicates that a given node has a relationship with other valuable nodes. A high eigenvector value for an actor means that a node has several neighbors with high eigenvector values
Eccentricity	The eccentricity of node v is calculated by computing the shortest path between node v and all other nodes in the graph; then the longest shortest path is chosen
Authority	Exhibits a node pointed to by many good hubs
Hub	Exhibits a node that points to many good authorities
PageRank	Assigns a numerical weight to each element of a hyperlinked set of documents, such as the World Wide Web, with the purpose of “measuring” its relative importance within the set
Clustering coefficient	It quantifies how close neighbors are to being a clique: a clique is a subset of all of the edges connecting pairs of vertices of an undirected graph

Table 2: The features belonging to the Network Measures category with their description.

aimed at verifying whether the behavioural features are able to discriminate the game bot attacks by the human user behaviour. The classification is carried out by using several state-of-the-art machine learning classifiers built with the behavioural feature categories we considered. The dataset involved in the study was obtained from the operation of Aion, a popular game and it is freely available. The dataset contains all in-game action logs for 88 days, between April 9th and July 5th of 2010. During this period, there were 49,739 players that played more than 3 h. Among these players, 7702 characters were game bots, identified by the game company. The banned list was provided by the game company to serve as the ground truth, and each banned user has been vetted and verified by human labor and active monitoring. With respect to log released by the game company, we aggregated the values of the features related to the same user in our dataset and we marked the feature vectors as “human” or “game bot” according to the game company indications. In order to assure the privacy of users the dataset is anonymized from user private and personal information. Indeed, the consent of users is taken into account by ensuring that data analysis is within the scope of end user license agreement: as a matter of fact, when users joined the Aion game, users granted to NCSOFT, Inc. the permission to use and share user data for analysis purpose.

The evaluation consists of three stages: (i) the comparison of descriptive statistics of the populations of behavioural features; (ii) the hypotheses testing, to verify if the feature categories have different distributions for the populations of game bot and human behaviour; and (iii) the classification analysis aimed at assessing whether the behavioural feature categories are able to correctly classify game bot and human behaviour. Relating to the descriptive statistics, we report the box plot of the distribution of game bot and human behaviour in order to demonstrate that the distributions are different. With regards to the hypotheses testing, the null hypothesis to be tested is:

> H_0: There are no statistically significant differences between the considered features of game bots and human users

The null hypothesis was tested with Mann-Whitney (with the p-level fixed to 0.05) and with Kolmogorov-Smirnov Test (with the p-level fixed to 0.05). We chose to run two different tests in order to enforce the conclusion validity. The purpose of these tests is to determine the level of significance, i.e., the risk (the probability) that erroneous conclusions be drawn: we set in following study the significance level equal to .05, i.e. we accept to make mistakes 5 times out of 100. The aim of the classification analysis is to assess if the features are able to correctly discriminate between game bot and human behaviours. We consider six different algorithms of classification: J48, DecisionStump, HoeffdingTree, RandomForest, RandomTree and REPTree (Canfora et al., 2015; Canfora et al., 2013). These algorithms were applied to the five features categories considered in the study (i.e., PI, PA, GA, SID and NM).

The classification analysis was performed using the Weka tool[4], a well-known suite of machine learning software.

4.1 Descriptive statistics

Figures 1, 2, 3, 4, 5, 6, 7, 8 and 9 show the box plots for a subset of features belonging to each category considered in the study. For reasons space, we do not show the box plots related to all features, but they lead to the same considerations.

Figure 1 shows the box plots related to the login frequency time feature (PI_1) between game bot and human users. The box plots are very similar, we conclude that for the feature point of view game bot and human user login with similar frequency. In our opinion the explanation of the obtained result is that the feature does not consider the login time but only the login frequency.

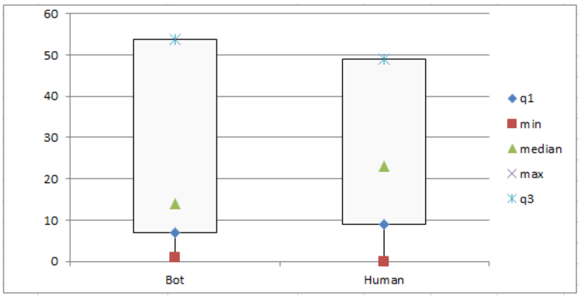

Figure 1: The box plot relating to the game bot and human distributions for the login frequency feature, belonging to the Player Information category.

The box plot in Figure 2 is related to the play time feature (PI_2). In this box plot the distributions between game bots and human users are very different: as matter of fact, while human user presents a small distribution, the game bot one exhibits wider values if compared with the human user one. The reason why this happens is that game bot are able to play the game without the need to break, differently from human users.

Figure 3 shows the box plot related to game bot and human distributions for sitting features (PA_1). The sitting feature is related to the number of rounds that game bots and human users are able to play. This result is consistent with the one we obtained by analyzing the previous box plot: as matter of fact game bots distribution seems to be wider if compared with the human users one.

[4]http://www.cs.waikato.ac.nz/ml/weka/

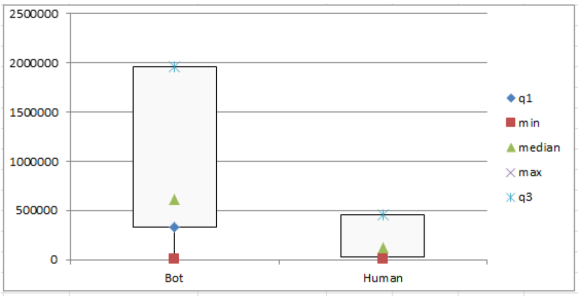

Figure 2: The box plot relating to the game bot and human distributions for the play time feature, belonging to the Player Information category.

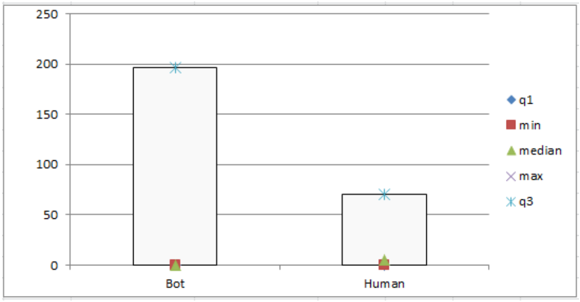

Figure 3: The box plot relating to the game bot and human distributions for the sitting feature, belonging to the Player Actions category.

Figure 4 shows the box plot related to earning experience points feature (PA_2). This feature is related to the ability to earn points in order to buy power for the character. Confirming what we expected, game bots are able to gather more points if compared with human users: this results is reflected in the wider dimension of game bots distribution with respect with the human users one.

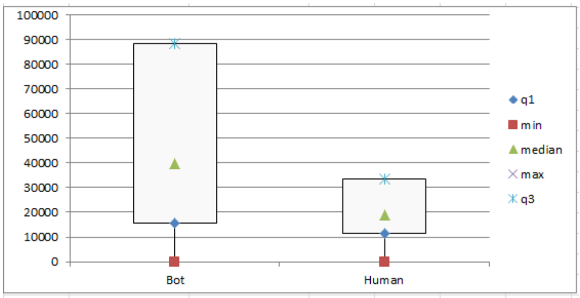

Figure 4: The box plot relating to the game bot and human distributions for the earning experience points feature, belonging to the Player Actions category.

Figure 5 shows the box plot related to the party play time feature (GA_1). In this case, the human user box plots are wider if compared with the game bots one. This happens because usually game bots have not interest to make party in order to play with other users: the focus of game bots is only to disturb human users and gather points in order to have character reinforcements.

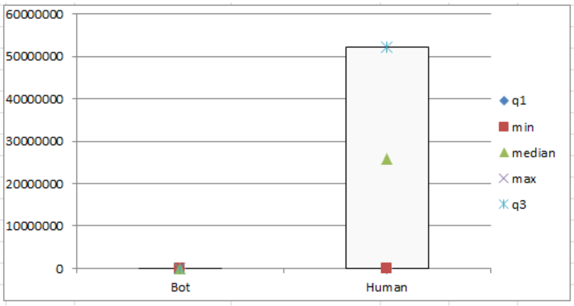

Figure 5: The box plot relating to the game bot and human distributions for the party play time feature, belonging to the Group Activities category.

Figure 6 shows the box plots related to the party play time feature (GA_2). The feature is related to the capacity of player to perform role missions. Confirming the previous box plot, game bot have no interest to cooperate in order to play, while human user usually need to play together in order to complete successfully difficult missions.

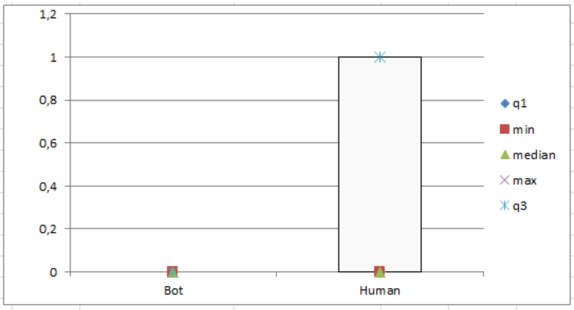

Figure 6: The box plot relating to the game bot and human distributions for the guild activities feature, belonging to the Group Activities category.

Figure 7 show the distributions for the SID_1 feature. The two distributions appear to be very similar: this is symptomatic of the fact the game bots like human users interact between them, for this reason the considered feature is not discriminative between the two distributions.

Figure 8 shows the box plot related to the NM_1 feature. These box plots exhibit that human users presents a wider degree centrality if compared with game bots. We conclude from these box plots that human users hare more friends if compared with game bots, for this reason the human users distribution is wider if compared with game bots one.

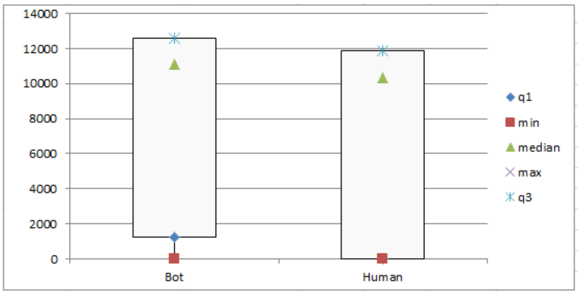

Figure 7: The box plot relating to the game bot and human distributions for the party play, belonging to the Social Interaction category.

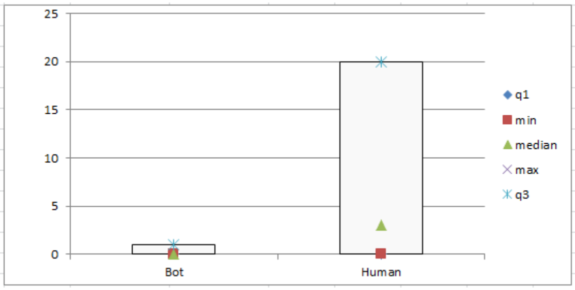

Figure 8: The box plot relating to the game bot and human distributions for the degree centrality feature, belonging to the Network Measures category.

Figure 9 shows the NM_2 feature box plots. As in the previous box plots, the human users distribution is wider than the game bots one. It is happen because in order to reach an objective or another player, the human user try to reach it using the shortest path, differently game bots do not consider this: this is reflected by the different distributions of the box plots.

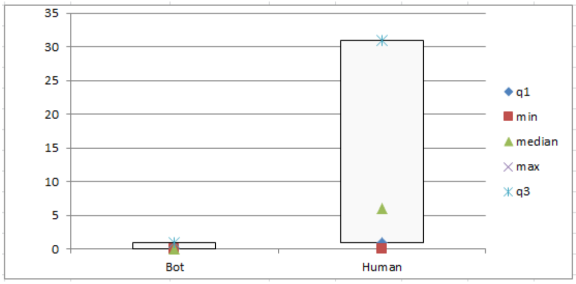

Figure 9: The box plot relating to the game bot and human distributions for the betweenness centrality feature, belonging to the Network Measures category.

4.2 Hypothesis testing

The hypothesis testing aims at evaluating if the features present different distributions for the populations of game bot and human behavioural characteristics with statistical evidence.

We assume valid the results when the null hypothesis is rejected by both the tests performed.

Table 7 shows the results of hypothesis testing: the null hypothesis H_0 can be rejected for all the features. This means that there is statistical evidence that the feature vector is a potential candidate for correctly classifying between game bot and human behavioural characteristics.

Variable	Mann-Whitney	Kolmogorov-Smirnov
PI1	0,000000	p < .001
PI2	0,000000	p < .001
PI3	0,000000	p < .001
PI4	0,000000	p < .001

Table 3: Results of the test of the null hypothesis H_0

Variable	Mann-Whitney	Kolmogorov-Smirnov
PA1	0,000000	p < .001
PA2	0,000000	p < .001
PA3	0,000000	p < .001
PA4	0,000000	p < .001
PA5	0,000000	p < .001
PA6	0,000000	p < .001
PA7	0,000000	p < .001
PA8	0,000000	p < .001
PA9	0,000000	p < .001
PA10	0,000000	p < .001

Table 4: Results of the test of the null hypothesis H_0

Variable	Mann-Whitney	Kolmogorov-Smirnov
GA1	0,000000	p < .001
GA2	0,000000	p < .001

Table 5: Results of the test of the null hypothesis H_0

This result will provide an evaluation of the risk to generalize the fact that the selected features produce values which belong to two different distributions (i.e., the one related to the game bots and the human users): those features can distinguish those observations. With the classification analysis we will be able to establish the accuracy of the features in associating any behavioural feature to the game bot or to the human distribution.

4.3 Classification analysis

Four metrics were used to evaluate the classification results: Precision, Recall, F-Measure and ROC Area.

The precision has been computed as the proportion of the examples that truly belong to class X among all those which were assigned to the class. It is the ratio of the number of relevant records retrieved

Variable	Mann-Whitney	Kolmogorov-Smirnov
SI1	0,000000	p < .001

Table 6: Results of the test of the null hypothesis H_0

Variable	Mann-Whitney	Kolmogorov-Smirnov
NM1	0,000000	p < .001
NM2	0,000000	p < .001
NM3	0,000000	p < .001
NM4	0,000000	p < .001
NM5	0,000000	p < .001
NM6	0,000000	p < .001
NM7	0,000000	p < .001
NM8	0,000000	p < .001
NM9	0,000000	p < .001

Table 7: Results of the test of the null hypothesis H_0

to the total number of irrelevant and relevant records retrieved:

$$Precision = \frac{tp}{tp+fp}$$

where *tp* indicates the number of true positives and *fp* indicates the number of false positives.

The recall has been computed as the proportion of examples that were assigned to class X, among all the examples that truly belong to the class, i.e., how much part of the class was captured. It is the ratio of the number of relevant records retrieved to the total number of relevant records:

$$Recall = \frac{tp}{tp+fn}$$

where *tp* indicates the number of true positives and *fn* indicates the number of false negatives.

The F-Measure is a measure of a test's accuracy. This score can be interpreted as a weighted average of the precision and recall:

$$F\text{-}Measure = 2 * \frac{Precision*Recall}{Precision+Recall}$$

The Roc Area is defined as the probability that a positive instance randomly chosen is classified above a negative randomly chosen.

The classification analysis consisted of building classifiers in order to evaluate the feature vector accuracy to distinguish between game bots and human users.

For training the classifier, we defined T as a set of labeled traces *(M, l)*, where each M is associated to a label $l \in \{B, H\}$ (where B represents the game bot, while H the human user). For each M we built a feature vector F $\in R_y$, where y is the number of the features used in training phase ($y = 4$ for the PI categoty, $y = 9$ for the PA category, $y = 2$ for the GA category, $y = 1$ for the SID category and $y = 9$ for the NM category).

For the learning phase, we use a k-fold cross-validation: the dataset is randomly partitioned into k subsets. A single subset is retained as the validation dataset for testing the model, while the remaining $k-1$ subsets of the original dataset are used as training data. We repeated the process for $k = 10$ times; each one of the k subsets has been used once as the validation dataset. To obtain a single estimate, we computed the average of the k results from the folds.

We evaluated the effectiveness of the classification method with the following procedure:

1. build a training set $T \subset D$;
2. build a testing set $T' = D \div$T;
3. run the training phase on T;
4. apply the learned classifier to each element of T'.

Each classification was performed using 20% of the dataset as training dataset and 80% as testing dataset employing the full feature set.

The results that we obtained with this procedure are shown in table 8.

We obtain following best results for each feature category we consider:

- a precision equal to 0.952 and a recall a 0.953 using the J48 algorithm classifying the feature belonging to PI category;
- a precision equal to 0.954 and a recall a 0.955 using the RandomForest algorithm classifying the feature belonging to PA category;
- a precision equal to 0.858 and a recall a 0.882 using the RepTree algorithm classifying the feature belonging to GA category;
- a precision equal to 0.836 and a recall to 0.875 using the J48 algorithm classifying the feature belonging to SID category;
- a precision equal to 0.923 and a recall 0.928 using the RandomForest algorithm classifying the feature belonging to NM category.

Considering that we obtained the best results in terms of precision and recall when classifying the features related to the PI and PA categories, we perform a feature selection on these categories in order to investigate whether using a small feature set we are able to obtain better results. As matter of fact, the feature selection is employed to improve the ability of classifier in discriminating human and game bot instances and decrease training time. We use as feature selection algorithm the BestFirst one, that implements a a best-first search strategy to navigate attribute subsets which basically explores a graph by expanding the most promising node chosen according to a specified rule.

Results of the feature selection are shown in Table 9.

We obtained that the most discriminating feature in the PI category is just the play time one, while in PA category are four features: sitting, earning experience points, obtaining items and earning player kill points.

In order to evaluate whether the two new feature set belonging to PI and PA categories are able to overcome the previous classifiers we learned, we build different classifiers with the features resulting from the features selection step: Table 10 shows the results we obtained.

The value of precision and recall are increased for both the features categories we considered:

- relating to the PI category, the best precision value it is incremented from 0.952 to 0.954 and the recall one is incremented from 0.953 to 0.984;
- relating to the PA category, the best precision value it is incremented from 0.954 to 0.960 and the recall one is incremented from 0.955 to 0.986;

The time analysis confirms the effectiveness of the feature selection step in order to quickly identify the game bot: as matter of fact with the exception of the RandomForest algorithm, the remaining ones are able to learn the classifiers in less than 1 second: for this reason the developed classifiers are able to discriminate the game bot from the human user in less than one second in the worst case (i.e., when the classifier is built with the new data).

RQ response: after the feature step selection, the best feature set is represents by following feature belonging to PA category: sitting, earning experience points, obtaining items and earning player kill points. Classifying with this feature set we obtain, using the RandomForest classification algorithm, a precision equal to 0.96 and a recall equal to 0.986.

5 CONCLUSIONS

The online game is a widespread form of entertainment. Through online worlds, users can play with other players, win upgrades for their character by defeating monsters and enemies, and even sell characters boosted by buying them through the game currency that can be changed in the real world one. In this scenario, the so-called malicious game bot, software able to purchase points without any interruption disturbing the game of human players, are widespread. Indeed attackers usually use game bots in order to obtain personal gain by selling its enhanced characters. In this paper we propose a method able to

Category	Algorithm	Precision	Recall	F-Measure	Roc Area	Time
PI	J48	0.95	0.951	0.949	0.856	1.97
	DecisionStump	0,942	0,944	0,941	0,823	0.21
	HoeffdingTree	0,941	0,944	0,941	0,872	0.3
	RandomForest	0,952	0,953	0,950	0,895	37.8
	RandomTree	0,917	0,916	0,916	0,814	0.64
	REPTree	0,946	0,948	0,945	0,880	0.93
PA	J48	0,948	0,950	0,947	0,862	7.01
	DecisionStump	0,934	0,936	0,935	0,830	0.58
	HoeffdingTree	0,942	0,944	0,942	0,872	0.94
	RandomForest	0,954	0,955	0.953	0,95	57.23
	REPTree	0,948	0,950	0,948	0,888	3.36
GA	J48	0,858	0,881	0,843	0,764	0.38
	DecisionStump	0,764	0,874	0,816	0,721	0.05
	HoeffdingTree	0,854	0,880	0,839	0,783	0.17
	RandomForest	0,820	0,855	0,833	0,766	17.71
	RandomTree	0,820	0,854	0,833	0,758	0.42
	REPTree	0,858	0,882	0,845	0,784	0.47
SID	J48	0,836	0,875	0,821	0,698	0.37
	DecisionStump	0,764	0,874	0,816	0,690	0.06
	HoeffdingTree	0,826	0,874	0,823	0,710	0.32
	RandomForest	0,805	0,863	0,822	0,688	8.31
	RandomTree	0,805	0,865	0,822	0,682	0.14
	REPTree	0,829	0,874	0,823	0,717	0.11
NM	J48	0,917	0,923	0,917	0,810	24.75
	DecisionStump	0,871	0,884	0,875	0,673	0.99
	HoeffdingTree	0,892	0,903	0,891	0,769	2.68
	RandomForest	0,923	0,928	0,923	0,858	42.42
	RandomTree	0,886	0,887	0,887	0,751	0.69
	REPTree	0,917	0,923	0,917	0,845	5.5

Table 8: Classification results: Precision, Recall, F-Measure and RocArea for classifying the feature categories, computed with six different classification algorithms. The Time column represents the time in seconds taken to build the model.

Table 9: Feature Selection Results.

#	Feature
PI	*play time*
PA	*sitting*
PA	*earning experience points*
PA	*obtaining items*
PA	*earning player kill points*

discriminate an human user from a game bot classifying a set of behavioural features. We consider features related to the player and related to the game, obtaining in the best case a precision equal to 0.96 and a recall equal to 0.986.

As future work we plan to investigate whether the feature vector we considered in this work is able to detect bot in social network. Furthermore, we will consider the adoption of Process Mining techniques (Bernardi et al., 2016) and Formal Methods (Francesco et al., 2016) in order to extract the game bots patterns with the aim to verify whether are different from the human users one.

ACKNOWLEDGEMENTS

This work has been partially supported by H2020 EU-funded projects NeCS and C3ISP and EIT-Digital Project HII.

Category	Algorithm	Precision	Recall	F-Measure	Roc Area	Time
PI_selected	J48	0,954	0,984	0,969	0,824	0.36
	DecisionStump	0,954	0,984	0,969	0,823	0.05
	HoeffdingTree	0,953	0,985	0,968	0.845	0.41
	RandomForest	0,943	0,944	0,943	0.831	15.06
	RandomTree	0,943	0,944	0,943	0.774	0.37
	REPTree	0,954	0,984	0,969	0.837	0.33
PA_selected	J48	0,958	0,986	0,972	0.870	0.48
	DecisionStump	0,957	0,971	0,964	0.830	0.15
	HoeffdingTree	0,957	0,985	0,971	0.882	0.25
	RandomForest	0,960	0,986	0,973	0.890	56.67
	RandomTree	0,954	0,954	0,954	0.818	0.98
	REPTree	0,959	0,985	0,972	0.889	0.74

Table 10: Feature Selection Classification results: Precision, Recall, F-Measure and RocArea for classifying the feature resulting of the feature selection process with the features of PI and PA categories, computed with six different classification algorithms. The Time column represents the time in seconds taken to build the model.

REFERENCES

Adams, E. (2014). *Fundamentals of game design*. Pearson Education.

Bernardi, M. L., Cimitile, M., Di Francescomarino, C., and Maggi, F. M. (2016). Do activity lifecycles affect the validity of a business rule in a business process? *Inf. Syst.*, 62(C):42–59.

Bernardi, M. L., Cimitile, M., Di Lucca, G. A., and Maggi, F. M. (2012). M3d: A tool for the model driven development of web applications. In *Proceedings of the Twelfth International Workshop on Web Information and Data Management*, WIDM '12, pages 73–80, New York, NY, USA. ACM.

Canfora, G., De Lorenzo, A., Medvet, E., Mercaldo, F., and Visaggio, C. A. (2015). Effectiveness of opcode ngrams for detection of multi family android malware. In *Availability, Reliability and Security (ARES), 2015 10th International Conference on*, pages 333–340. IEEE.

Canfora, G., Mercaldo, F., and Visaggio, C. A. (2013). A classifier of malicious android applications. In *Availability, Reliability and Security (ARES), 2013 Eighth International Conference on*, pages 607–614. IEEE.

Chen, K.-T., Pao, H.-K. K., and Chang, H.-C. (2008). Game bot identification based on manifold learning. In *Proceedings of the 7th ACM SIGCOMM Workshop on Network and System Support for Games*, pages 21–26. ACM.

Chen, Y.-C., Chen, P. S., Song, R., and Korba, L. (2004). Online gaming crime and security issue-cases and countermeasures from taiwan. In *PST*, pages 131–136.

Chung, Y., Park, C. Y., Kim, N., Cho, H., Yoon, T. B., Lee, H., and Lee, J. (2015). A behavior analysis-based game bot detection approach considering various play styles. *CoRR*, abs/1509.02458.

Francesco, N. D., Lettieri, G., Santone, A., and Vaglini, G. (2016). Heuristic search for equivalence checking. *Software and System Modeling*, 15(2):513–530.

Griffiths, M. D., Davies, M. N., and Chappell, D. (2004). Online computer gaming: a comparison of adolescent and adult gamers. *Journal of adolescence*, 27(1):87–96.

Hilaire, S., Kim, H.-c., and Kim, C.-k. (2010). How to deal with bot scum in mmorpgs? In *Communications Quality and Reliability (CQR), 2010 IEEE International Workshop Technical Committee on*, pages 1–6. IEEE.

Kang, A. R., Jeong, S. H., Mohaisen, A., and Kim, H. K. (2016). Multimodal game bot detection using user behavioral characteristics. *SpringerPlus*, 5(1):523.

Kashifuji, Y. (2008). Detection of mmorpg bots based on behavior analysis. *ACE*, 2008:4.

Kim, H., Hong, S., and Kim, J. (2005). Detection of auto programs for mmorpgs. In *Australasian Joint Conference on Artificial Intelligence*, pages 1281–1284. Springer.

Mishima, Y., Fukuda, K., and Esaki, H. (2013). An analysis of players and bots behaviors in mmorpg. In *Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on*, pages 870–876. IEEE.

Oh, J., Borbora, Z. H., Sharma, D., and Srivastava, J. (2013). Bot detection based on social interactions in mmorpgs. In *Social Computing (SocialCom), 2013 International Conference On*, pages 536–543. IEEE.

Paulson, R. A. and Weber, J. E. (2006). Cyberextortion: an overview of distributed denial of service attacks against online gaming companies. *Issues in Information Systems*, 7(2):52–56.

Quandt, T. and Kröger, S. (2013). *Multiplayer: The social aspects of digital gaming*, volume 3. Routledge.

Seay, A. F., Jerome, W. J., Lee, K. S., and Kraut, R. E. (2004). Project massive: a study of online gaming communities. In *CHI'04 extended abstracts on Human factors in computing systems*, pages 1421–1424. ACM.

Thawonmas, R., Kashifuji, Y., and Chen, K.-T. (2008). Detection of mmorpg bots based on behavior analysis. In *Proceedings of the 2008 International Conference* *on Advances in Computer Entertainment Technology*, pages 91–94. ACM.

Wellman, B. and Gulia, M. (1999). Virtual communities as communities. *Communities in cyberspace*, pages 167–194.

Yampolskiy, R. V. and Govindaraju, V. (2008). Embedded noninteractive continuous bot detection. *Computers in Entertainment (CIE)*, 5(4):7.

Yee, N. (2008). Maps of digital desires: Exploring the topography of gender and play in online games. *Beyond Barbie and Mortal Kombat: New perspectives on gender and gaming*, pages 83–96.