

Game-bot Detection based on Analysis of Harvest Coordinate

Jae Woong Choi*, Ah Reum Kang**

*Student, Dept. of Information Security, Korea University, Seoul, Korea

**Professor, Dept. of Information Security, Pai Chai University, Daejeon, Korea

[Abstract]

As the online game market grows, the use of game bots is causing the most serious problem for game services. We propose a harvest coordinate analysis model to detect harvesting bots among game bots of the Massively Multiplayer Online Role-Playing Games(MMORPGs) genre. The proposed model analyzes the player's harvesting behavior using the coordinate data. Game bots can obtain in-game goods and items more easily than normal players and are not affected by realistic restrictions such as sleep time and character manipulation fatigue. As a result, there is a difference in harvesting coordinates between normal players and game bots. We divided the coordinate zones and used these coordinate zone differences to distinguish between game bot players and normal players. We created a dataset with NCSOFT's AION log and applied it to a random forest model to detect game bots, and as a result, we derived performance with a recall of 0.72 and a precision of 0.92.

▶ **Key words:** online game security, game bot, coordinate analysis, MMORPG, behavior analysis

[요 약]

온라인 게임 시장이 성장하면서 게임 봇의 사용은 게임 서비스에 가장 심각한 문제를 야기하고 있다. 본 논문에서는 MMORPG 장르의 게임 봇 중 채집을 진행하는 봇을 탐지하기 위한 채집 좌표 분석 모델을 제안한다. 제안한 모델은 좌표 데이터를 기반으로 플레이어의 채집 행위를 분석한다. 정상적인 플레이어보다 손쉽게 게임 내 재화와 아이템을 수급할 수 있는 게임 봇은 수면 시간, 캐릭터 조작 피로도와 같은 현실적인 제약의 영향을 받지 않기 때문에 채집 행위를 시도하는 좌표 구역에 차이가 발생한다. 좌표 구역을 나누고 각 플레이어의 좌표 구역 차이를 이용하여 게임 봇 플레이어와 정상적인 플레이어를 구분해 낼 수 있도록 했다. NCSOFT사의 AION 로그로 데이터셋을 만들고 random forest 모델에 적용하여 게임 봇을 탐지한 결과 재현율 72%, 정밀도 92%의 성능을 보였다.

▶ **주제어:** 온라인 게임 보안, 게임 봇, 좌표 분석, MMORPG, 행위 분석

• First Author: Jae Woong Choi, Corresponding Author: Ah Reum Kang
*Jae Woong Choi (cjbizwoong@korea.ac.kr), Dept. of Information Security, Korea University
**Ah Reum Kang (armk@pcu.ac.kr), Dept. of Information Security, Pai Chai University
• Received: 2022. 03. 18, Revised: 2022. 04. 25, Accepted: 2022. 04. 25.

I. Introduction

온라인 게임 시장의 발전과 동시에 온라인 게임 보안은 중요한 이슈가 되었다. 온라인 게임 보안에는 다양한 목표가 있으나 가장 중요하게 여겨지는 보안 이슈는 온라인 게임 시장과 함께 발전해 나가는 게임 봇의 확산이다. 온라인 게임에서는 다양한 경로를 통해 인게임 재화를 획득할 수 있으며, 플레이어들은 인게임 재화를 현실 재화로 환전할 수 있다. 플레이어 간의 인게임 재화의 거래가 활발해지면서 일부 플레이어들은 더 많은 현실 재화를 얻기 위해 인게임 재화를 손쉽게 벌고자 했고, 게임 봇을 사용하여 악성 플레이를 수행하는 경우가 발생했다.

게임 봇은 정상 플레이어에게 존재하는 수면 시간과 같은 현실적인 제약이 존재하지 않고, 반복적인 작업을 자동으로 수행할 수 있기에 정상 플레이어보다 더욱 손쉽게 많은 인게임 재화를 수급할 수 있다. 이러한 게임 봇의 무분별한 재화 수집과 재화의 유통은 게임 내 시장에 인플레이션을 초래할 수 있으며, 정상 플레이어는 이에 대해 상대적인 박탈감을 느낄 수 있다. 그 결과 정상 플레이어는 게임을 이탈하거나 게임 봇을 사용하는 또 다른 악성 플레이어가 되는 경우도 존재한다. 이러한 문제는 게임 내 생태계를 더욱 파괴하게 되고 결과적으로 온라인 게임의 수명이 축소되는 원인이 된다.

MMORPG 장르에 존재하는 게임 봇의 종류에는 채집 봇, 사냥 봇이 가장 유명하며 사냥 봇의 경우 오픈 월드 사냥터에서 파티 플레이를 통해 반복 사냥을 수행하는 봇과 인스턴스 던전을 클리어하여 얻을 수 있는 아이템과 재화를 목적으로 파티 플레이를 수행하는 봇으로 세분화할 수 있다. 다양한 종류의 봇을 모두 탐지하기에는 실질적인 어려움이 존재하며 정상 플레이어와 구분이 되지 않는 특성을 가지는 경우도 존재한다.

본 연구에서는 플레이어의 행위 좌표 분석을 기반으로 한 게임 봇 탐지 모델을 제안한다. 2장에서는 기존의 연구를 검토하고 제안하는 모델과의 차이점을 알아보았다. 3장에서는 플레이어의 행위 좌표 분석을 기반으로 한 게임 봇 탐지 모델을 제안했다. 4장에서는 실제 서비스 중인 게임의 로그 데이터를 사용하여 제안한 모델을 평가했다. 5장에서는 결론과 앞으로의 연구 방향에 대해서 설명했다.

II. Preliminaries

1. Related works

과거부터 다양한 범주에서 다양한 특성을 가진 게임 봇을 탐지하는 연구가 진행되었다. Table 1.에 선행 연구를 두 가지 범주로 분류하고 정리했다.

첫 번째 범주는 플레이어와 게임 봇 사이에 행위 기반 탐지이다. 해당 범주에는 다음과 같은 선행 연구가 포함된다. Kang 등[1], [3]은 플레이어 간의 파티 플레이에 대한 로그를 기반으로 봇의 탐지를 진행하는 연구를 진행했다. Kang 등[2]은 플레이어들의 채팅 패턴을 분석하여 일반 플레이어와 게임 봇이 진행하는 채팅의 차이를 토대로 봇을 탐지하는 연구를 진행했다. Chung 등[5]은 봇의 플레이 스타일이 다르다는 가정을 통해 행동 유사성을 그룹화하여 게임 봇을 탐지하는 연구를 진행했다. Lee 등[6]은 플레이어의 전체 행동 시퀀스를 빅데이터 분석 플랫폼을 활용하여 게임 봇을 탐지하는 연구를 진행했다. Thawonmas 등[7]은 플레이어의 행위를 기반으로 임계값을 기반으로 봇을 분류하고 학습을 진행하여 게임 봇을 탐지하는 연구를 진행했다. Berandi 등[8]은 시계열 분석 기법을 활용하여 분석하고 MLP 신경 모델을 토대로 게임 봇을 탐지하는 연구를 진행했다. Kim과 Kim[9]은 게임 내 그룹 시스템인 길드 콘텐츠를 활용하여 길드 구성원의 행위를 분석하고 게임 봇의 그룹인 작업장을 탐지하는 연구를 진행했다. Song과 Kim[11]은 플레이어의 게임 내 재화 변동을 기반으로 클러스터링을 진행하여 게임 봇을 탐지하는 연구를 진행했다.

두 번째 범주는 플레이어와 게임 봇 사이의 사회적 상호작용의 차이가 있는 것을 토대로 분석하는 소셜 네트워크 기반 탐지이다. 해당 범주에는 다음과 같은 선행 연구가 포함된다. Jeong 등[4]은 플레이어를 그룹화하고 그룹별 사회적 상호작용 네트워크를 관찰하는 것으로 게임 봇의 행동을 분석하는 연구를 진행했다. Valvello와 Voelker[10]은 플레이어 캐릭터의 상호작용을 기반으로 소셜 그래프를 작성하고 분석하는 것으로 소셜 네트워크 기반의 봇 탐지 전략을 제시하고 있다. Fujita 등[12]은 게임 내 플레이어 간 거래 네트워크와 거래 데이터를 사용하여 게임 봇을 탐지하는 연구를 진행했다. Lee 등[13]은 여러 타입의 게임 봇을 네트워크 클러스터링을 활용해 분류하고 탐지하는 연구를 진행했다.

Table 1. Description of previous studies

| Category | Description | Methodology |
|----------------|---|---|
| user behavior | analysis methodology using difference of activity between player and bot | analysis based on party play log[1], [3] |
| | | analysis based on communication pattern[2] |
| | | analysis based on user behavioral similarity[5] |
| | | entire action sequence analysis based on big data analysis platform[6] |
| | | analysis based on discrepancies in action frequencies and action types between normal and bot characters[7] |
| | | time series analysis based on user behavioral[8] |
| | | analysis based on guild system[9] |
| | | analysis based on clustering about asset change coordinate[11] |
| social network | analysis methodology using difference of social activity between player and bot | analysis based on social interaction network of normal and bot group[4], [10] |
| | | analysis based on trading network and trading data[12] |
| | | analysis based on network clustering about categories of bot types[13] |

기존의 연구에서 제한된 행위들을 선택하여 분석 범위를 제한하거나 전체 행위를 토대로 분석 범위를 확장하여 연구를 진행했으나 좌표 데이터를 토대로 연구를 진행한 경우는 자산 변동 좌표에 대한 경우 밖에 존재하지 않는 것을 발견했다. 채집 활동의 경우 채집물의 생성 및 재생성 좌표가 특정 지점으로 지정되어 있거나 특정 범위로 지정된다는 것을 다룬 연구는 없었다.

III. The Proposed Model

1. Theory

일반적으로 MMORPG 장르의 게임을 플레이하는 플레이어들은 다양한 액션 행위를 효율적으로 한다. 특히 정해진 목적을 위해 반복적인 행위를 해야 하는 경우 더 효율적으로 행동하게 된다. 채집의 경우가 이에 해당하며 플레이어들은 게임 맵 내에 산발적으로 생성되는 채집 아이템을 효율적으로 채집하기 위해 많은 사람에게 알려진 혹은 자신만의 경로를 따라 채집을 진행하고는 한다. 플레이어가 채집 행위를 이러한 방식으로 진행하는 이유는 캐릭터의 이동 거리를 최소화하는 것으로 플레이어의 캐릭터 컨

트롤에 대한 피로도를 줄이기 위한 것이다. 반대로 자동 채집을 위한 게임 봇을 이용하는 악성 플레이어의 경우 캐릭터 컨트롤에 대한 현실에서의 피로도를 신경 쓰지 않아도 되며, 이는 일반 캐릭터와 봇 캐릭터의 채집 위치의 차이가 발생하는 것으로 연결된다.

채집을 시도한 좌표를 추출하고 일정 범위로 나누어 여러 개의 채집 구역으로 나타내면, 각 구역 내 채집 횟수를 구할 수 있다. 이를 활용하여 플레이어의 게임 플레이 스타일에 따라 선호하는 채집 구역과 다양한 채집 패턴을 확인할 수 있다. 나아가 채집 구역 내에서 봇 플레이어와 일반 플레이어의 비율을 확인하고 일반 플레이어와 봇 플레이어가 각각 선호하는 채집 구역이 존재하는지 확인할 수 있다. 제안하는 모델의 처리 흐름은 Fig. 1.과 같다.

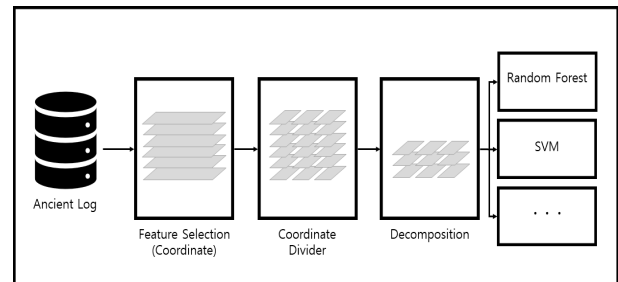


Fig. 1. Flow Chart about Proposed Model

2. Extracted Data

원시 로그는 NCSOFT 사의 게임인 AION의 로그를 활용했다. 전체 로그에서 아이템 채집과 아이템 채집 실패 로그를 추출하여 채집의 성공 및 실패와 관계없이 채집을 시도한 것에 대한 정보를 추출했다. 해당 로그에서 actor, actor_account, location_x, location_y, location_z 정보를 추출했다. actor와 actor_account 정보는 유저와 해당 유저의 계정을 식별하기 위한 정보이며, location 정보는 유저가 로그를 발생시킨 좌표 정보이다. 해당 로그의 사람 및 봇 플레이어에 대한 레이블은 actor_account 단위로 구성되어 있으며, 하나의 계정에서 불법 프로그램을 사용한 캐릭터는 정해져 있고, 다른 캐릭터는 사용하지 않아도 해당 유저의 계정을 제재하여 포함된 모든 캐릭터는 접속 불가능하다.

IV. Experiment and Evaluation

1. Dataset

본 연구에서 사용한 데이터셋은 원시 로그에서 추출한 채집 로그를 토대로 구성했다. 해당 데이터셋은 2010년 4월

16일부터 2010년 5월 7일까지 기간의 원시 로그 중 채집을 시도한 로그가 있는 플레이어의 정보를 추출했다. 데이터셋 내의 전체 플레이어 수는 11,840명이고 일반 플레이어로 분류된 수가 11,482명, 봇 플레이어로 분류된 수가 358명이다. 또한, 봇 플레이어로 분류된 캐릭터에는 채집을 수행하지 않고 사냥 및 거래 등의 매크로 행위만 반복하는 플레이어가 존재하기 때문에 채집 횟수의 이상치를 1,000으로 정하고 이상치 미만의 봇 플레이어를 학습에서 제외했다.

최종적으로 본 연구에서 사용한 데이터셋은 전체 플레이어 수가 11,740명으로 구성되어 있으며, 이 중 일반 플레이어로 분류된 수가 11,482명, 봇 플레이어로 분류된 수가 258명이다.

2. Feature Statistic Analysis

본 연구에서는 각 플레이어 별 전체 채집 횟수와 좌표 범위를 가공하여 3,677개의 feature를 만들었다. 3,677개의 feature는 3차원의 좌표 데이터를 일정 범위로 나누어 채집 구역으로 나타낸 각각의 좌표 범위이다. 전체 채집 횟수는 각 유저의 채집 구역과 관계없이 채집을 수행한 횟수이며, 다른 모든 feature는 나누어진 각각의 채집 구역 내에서 채집을 수행한 횟수이다.

3,677개의 채집 구역 특성은 전체 좌표 구역에서 채집이 발생하지 않는 구역을 제외하여 차원을 축소한 특성이다. 이러한 구역이 발생하는 이유는 게임의 모든 맵에 채집물이 생성되는 것은 아니며, 이동 불가 지역의 좌표도 존재하기 때문이다. Fig. 2. 는 x축, y축 및 z축을 일정 범위로 나눈 채집 구역 내에서 채집을 진행한 봇과 플레이어의 비율이다. 해당 파이 차트를 토대로 봇과 일반 플레이어가 서로 선호하는 채집 구역이 존재한다는 것을 확인할 수 있다. Fig. 3. 은 봇 3명과 일반 플레이어 3명의 채집 좌표를 산점도로 표현한 것이다. 각 축은 x, y, z 좌표이며 각 좌표의 최댓값으로 축 범위를 지정했다. 이를 통해 다시 일반 플레이어와 봇 플레이어는 어느 정도 선호하는 채집 구역이 나뉘어 있는 것을 확인할 수 있다. 이 외에도 많은 특성에서 채집 구역 내 플레이어의 비율이 고르거나 일반 혹은 봇 플레이어 중 하나로 편향된 결과가 나오는 것을 확인했다. 이와 같이 봇 플레이어와 일반 플레이어의 선호 채집 구역에 대한 차이가 발생하는 이유는 캐릭터 조작에 대한 피로도의 영향이 크다. 일반 플레이어의 경우 꾸준히 컴퓨터에 앉아서 채집을 수행하기에 채집과 같은 정적이고 반복적인 일을 하는 것에 피로감을 느끼게 된다. 그러나 봇 플레이어는 이러한 현실적인 제약이 존재하지 않기에 선호하는 채집 구역의 차이가 발생한다.

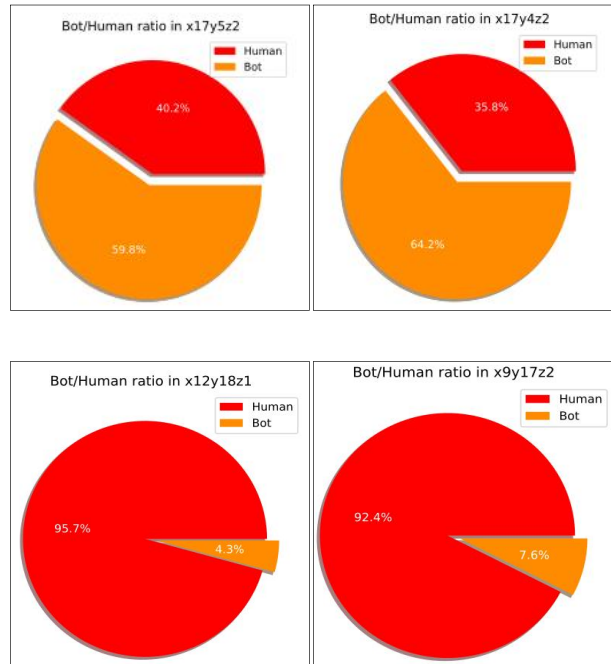


Fig. 2. Bot and human ratios in some harvesting area

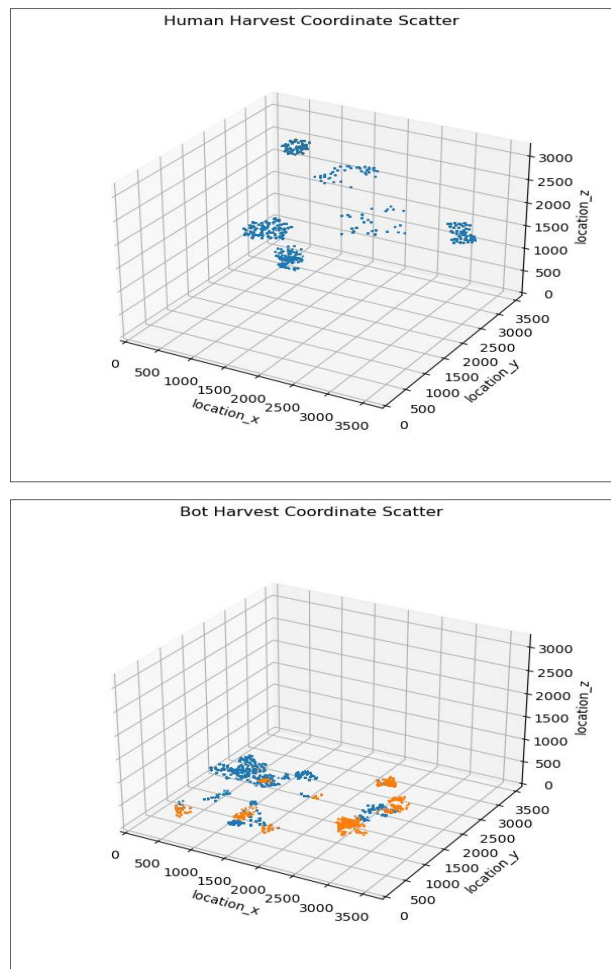


Fig. 3. Human & Bot Harvest Coordinate 3D Scatter

3. Machine Learning Algorithm

랜덤 포레스트 알고리즘은 결정 트리 알고리즘의 단점을 보완한 형태의 알고리즘으로 여러 개의 결정 트리 알고리즘을 학습하여 결과를 종합해 예측하는 앙상블 알고리즘이다. 성능이 우수하고 과적합 문제를 해결할 수 있으나, 결정 트리의 개수가 늘어나고 이들 간의 연관성이 개입하는 것으로 인해 분석이 어려워지는 경향이 존재한다. 그러나 좋은 성능으로 인해 다양하게 활용되는 알고리즘 중 하나이다.

Table 2. Performance

| Algorithm | Recall | Precision | F1-Score | Accuracy |
|-----------------------------------|--------|-----------|----------|----------|
| Logistic Regression[14] | 0.80 | 0.83 | 0.82 | 0.99 |
| Linear SVM[16] | 0.77 | 0.78 | 0.77 | 0.90 |
| Decision Tree[17] | 0.71 | 0.79 | 0.75 | 0.92 |
| Random Forest[15], [18] | 0.72 | 0.92 | 0.81 | 0.99 |
| AdaBoost[15] | 0.73 | 0.85 | 0.79 | 0.92 |
| Gradient Boost Machine [15], [19] | 0.65 | 0.87 | 0.75 | 0.98 |

Table 3. Performance by Each Ratio about Random Forest

| ratio (Train: Test) | Recall | Precision | F1-Score | Accuracy |
|---------------------|--------|-----------|----------|----------|
| 3:7 | 0.65 | 0.89 | 0.75 | 0.99 |
| 4:6 | 0.68 | 0.87 | 0.76 | 0.99 |
| 5:5 | 0.66 | 0.87 | 0.75 | 0.99 |
| 6:4 | 0.69 | 0.91 | 0.79 | 0.99 |
| 7:3 | 0.72 | 0.92 | 0.81 | 0.99 |

Table 4. Random Forest Confusion Matrix

| Actual \ Predict | Bot | Human |
|------------------|-----|-------|
| | Bot | 64 |
| Human | 5 | 3,429 |

본 연구에서는 다양한 머신러닝 알고리즘을 활용하여 실험을 진행했다. 총 11,740명의 플레이어를 활용해 학습 및 테스트를 진행했다. 기본적으로 학습과 테스트 비율은 7:3으로 진행했다. 각 머신 러닝 알고리즘 별 성능 평가 지표는 Table 2.와 같다. Random forest를 활용하여 학습을 진행하는 경우 선행 연구[1]보다 채집 로그를 활용한 봇 탐지에 더 높은 정확도를 보여주고 있는 것을 확인할 수 있다. 선행 연구[1]에서는 플레이어의 전체 행동을 기반으

로 분석을 진행하였으나 채집 로그를 활용한 경우 정확도가 54%로 낮은 수치를 보여주었다. 본 연구에서는 채집 로그를 활용하여 진행한 분석에서 99%의 정확도를 보여주고 있다. 본 논문에서는 실험 과정에서 도출된 평가 메트릭 중 precision이 가장 높은 random forest 알고리즘을 제안한다. Random forest 알고리즘의 학습과 테스트 비율을 조절하며 성능을 비교하였고, 그 결과는 Table 3.과 같다. 본 논문에서는 precision이 가장 높은 train과 test 비율이 7:3인 모델을 제안한다. 해당 모델에 대한 confusion matrix는 Table 4.와 같다. Confusion matrix를 확인하면 데이터 불균형으로 인해 정확도가 높은 것을 알 수 있다. 그럼에도 불구하고 random forest 알고리즘은 precision이 높은 수치를 보여주고 있으며, 이로 인해 정상 플레이어를 봇 플레이어로 오탐하는 경우를 줄일 수 있게 실제 서비스에 적합한 결과를 확인할 수 있다.

선행 연구[1]보다 더 좋은 결과가 나온 이유는 분석 범위의 차이 때문이다. 선행 연구[1]의 경우 전체 봇을 탐지하기 위해 분석 범위를 넓혔으나, 본 논문에서 제시하는 모델은 채집 행위에 대한 봇으로 분석 범위를 좁혔기에 채집 precision이 높게 나왔다.

4. Feature Importance Analysis

Random forest를 활용하여 학습을 진행한 후 각 특성의 중요도를 추출했다. 추출한 특성의 중요도의 상위 10개는 Fig. 4.와 같다. 본 연구에서 사용한 데이터셋의 경우 특성이 3,678개이지만 3,677개의 특성이 좌표 데이터를 활용하여 채집 구역을 나타낸 것이기 때문에 각각의 중요도는 낮게 측정이 되었으나 전체 학습의 97%의 영향력을 미치고 있는 것을 Fig. 4.에서 확인할 수 있다. 반대로 채집 횟수(harv_cnt)를 나타내는 특성의 경우 개별적으로 3%의 영향력을 미치고 있으며, 3,677개의 채집 구역 특성 중 특정 특성은 다른 채집 구역보다 더 많은 영향력을 미치고 있는 것을 확인했다. 해당 그림의 y축 값으로 사용된 $x[n_1]y[n_2]z[n_3]$ 는 좌표 범위를 의미한다. 각각의 n_1, n_2, n_3 의 범위 표현식은 아래 수식과 같다.

$$pos(n_i) (i = 1, 2, 3) = (n_i + 1) * 100 - 100 \leq n_i < (n + 1) * 100 \quad (1)$$

사용한 데이터셋의 모든 feature는 채집 횟수를 나타내고 있다. 전체 채집 횟수에 대한 harv_cnt 특성은 전체 구역에서의 채집 횟수이기 때문에 가장 높은 특성 중요도를 가진다고 판단했다. 더 나아가 전체 채집 횟수로 구분되지 않는 플레이어의 경우 채집 구역별 수치를 토대로 구분했다. 정상 혹은 봇 플레이어별 선호 채집 구역이 존재하기 때문에 특정 구역에서 높은 특성 중요도를 나타낸다고 판단했다.

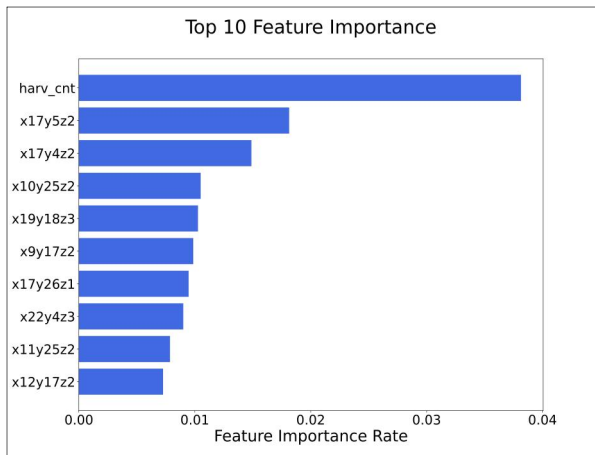


Fig. 4. Top 10 Feature Importance

V. Conclusions & Future Work

본 연구에서는 온라인 게임 봇 탐지를 위해 좌표 데이터를 사용하여 행위 기반 분석 모델을 제안한다. 제안하는 모델은 실제 서비스 중인 MMORPG 장르의 게임에 대부분 존재하는 채집 콘텐츠를 기반으로 좌표 데이터를 분석하여 특성 셋을 구성하기 때문에 다양한 MMORPG 게임에 적용할 수 있다.

게임 플레이어의 채집 시도 좌표를 토대로 게임 봇과 정상 플레이어별 선호 채집 구역을 사용했다. 채집을 수행하는 정상 플레이어들은 캐릭터 조작에 대한 피로도로 인해 채집 구역을 손쉽게 순회할 수 있는 구역을 찾아 편하게 채집을 진행하는 경향이 있다. 그러나 게임 봇의 경우 피로도를 느끼지 않으며, 현실적인 제약에 영향을 받지 않기 때문에 채집 아이템의 생성 및 재생성 시간에 맞추어 더욱 넓은 채집 구역을 순회할 수 있다. 따라서 게임 봇은 정상 플레이어보다 더 많은 양의 채집 아이템을 획득할 수 있게 된다. 이런 이유로 게임 봇과 정상 플레이어는 채집 구역 및 채집 패턴에 차이를 보인다.

NCSOFT사의 게임인 AION 로그에 제안한 모델을 적용하였다. 임의로 분할한 채집 구역을 시각화하여 게임 봇과 정상 플레이어 사이에 선호하는 채집 구역의 차이가 발생한다는 사실을 알게 되었다. 채집 구역별 채집 횟수를 feature로 사용하여 학습을 진행한 결과 random forest 알고리즘으로 72%의 재현율과 92%의 정밀도를 얻었다.

향후 연구에서는 채집 구역을 세밀하게 구분하고 클러스터링을 활용하여 좌표 데이터 기반 게임 봇 탐지를 진행해보고자 한다. 또한, 채집 이외에 좌표 데이터를 활용할 수 있는 다양한 플레이어 행위를 분석하여 어느 게임에서

나 서버 단에 기록되는 로그를 토대로 제안한 모델을 적용할 수 있게 할 것이다.

이어서 다양한 종류의 게임 봇을 특성별로 구분하고 학습한 모델을 모듈화할 수 있다면 분석 범위를 축소하여 좋은 성능을 내는 모델 여러 개를 하나의 프레임워크로 작성할 수 있을 것이라는 가설을 세웠고, 이를 토대로 더 나은 성능을 보여주는 프레임워크를 구성할 수 있을 것이다.

ACKNOWLEDGEMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2020R111A1A01061146).

REFERENCES

- [1] A.R. Kang, J. Woo, J. Park, and H.K. Kim, "User behavior analysis for online game bot detection," *Journal of The Korea Institute of Information Security & Cryptology*, Vol. 22, No. 2, pp. 225-238, Apr. 2012. DOI: 10.13089/JKIISC.2012.22.2.225
- [2] A.R. Kang, H.K. Kim, and J. Woo, "Chatting pattern based game BOT detection: do they talk like us?," *KSIIT Transactions on Internet and Information Systems (TIIS)*, Vol. 6, No. 11, pp. 2866-2879, Nov. 2012. DOI: 10.3837/tiis.2012.10.007
- [3] A.R. Kang, J. Woo, J. Park, and H.K. Kim, "Online game bot detection based on party-play log analysis," *Computers and Mathematics with Applications*, Vol. 65, No. 9, pp. 1384-1395, May. 2013. DOI: 10.1016/j.camwa.2012.01.034
- [4] S.H. Jeong, A.R. Kang, and H.K. Kim, "Analysis of Game Bot's Behavioral Characteristics in Social Interaction Networks of MMORPG," *ACM SIGCOMM Computer Communication Review*, Vol. 45, No. 4, pp. 99-100, Sep. 2015.
- [5] Y. Chung, C. Park, N. Kim, H. Cho, T. Yoon, H.J. Lee, and J.H. Lee, "Game Bot Detection Approach Based on Behavior Analysis and Consideration of various Play Styles," *ETRI Journal*, Vol. 35, No. 6, pp. 1058-1067, Dec. 2013. DOI: 10.4218/etrij.13.2013.0049
- [6] J. Lee, J. Lim, W. Cho, and H.K. Kim, "In-Game Action Sequence Analysis for Game Bot Detection on the Big Data Analysis Platform," *Proceedings of the 18th Asia Pacific Symposium on Intelligent and Evolutionary Systems*, Vol. 2, pp. 403-414, Jan. 2015.
- [7] R. Thawonmas, Y. Kashifuji, and K.T. Chen, "Detection of MMORPG Bots Based on Behavior Analysis," *Proceedings of the 2009 International Conference on Advances in Computer*

- Entertainment Technology, pp. 91-94, Dec. 2008.
- [9] H. Kim and H.K. Kim, "Research on online game bot guild detection method," *Journal of The Korea Institute of information Security & Cryptology*, Vol. 25, No. 5, pp. 1115-1122, Oct. 2015. DOI: 10.13089/JKIISC.2015.25.5.1115
- [10] M. Varvello and G.M. Voelker, "Second life: a social network of humans and bots," *Proceedings of the 20th international workshop on network and operating systems support for digital audio and video*, pp. 9-14, Jun. 2010.
- [11] H.M. Song and H.K. Kim, "Game-bot detection based on Clustering of asset-varied location coordinate," *Journal of The Korea Institute of information Security & Cryptology*, Vol. 25, No. 5, pp. 1131-1141, Oct. 2015. DOI: 10.13089/JKIISC.2015.25.5.1131
- [12] A. Fujita, H. Itsuki, and H. mastsubara, "Detecting Real Money Traders in MMORPG by Using Trading Network," *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 6, No. 1, pp. 26-31, Oct. 2011.
- [13] E. Lee, J. Lee, and J. Kim, "Detecting the bank character in MMORPGs by analysis of a clustered network," *The 3rd International Conference on Internet*, Dec. 2011.
- [14] M.P. LaValley, "Logistic regression," *Circulation*, Vol. 117, No. 18, pp. 2395-2399, May. 2008.
- [15] N.C. Oza, and S. J. Russell, "Online bagging and boosting," *International Workshop on Artificial Intelligence and Statistics*, PMLR, pp. 229-236, 2011.
- [16] A. Ukil, "Support vector machine," *Intelligent Systems and Signal Processing in Power Engineering*, Springer, Berlin, Heidelberg, pp. 161-226, 2007.
- [17] A. J. Myles, R.N. Feudale, Y. Liu, N.A. Woody, and S.D. Brown, "An introduction to decision tree modeling," *Journal of Chemometrics: A Journal of the Chemometrics Society*, Vol. 18, No. 6, pp. 275-285, Nov. 2004. DOI: 10.1002/cem.873
- [18] T.M. Oshiro, P.S. Perez, and J.A. Baranauskas, "How Many Trees in a Random Forest," *International workshop on machine learning and data mining in pattern recognition*, Springer, Berlin, Heidelberg, 2012.
- [19] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Frontiers in neurobotics*, 7: 21, Dec. 2013.

Authors



Jae Woong Choi received the B.S. degree in Cyber Security from Pai Chai University, Korea, in 2022. He is currently pursuing the M.S. degree in Information Security from Korea University. His research interest

includes online game security, block-chain, NFT game security.



Ah Reum Kang received the M.S. and Ph.D. degrees in Information Security from Korea University, Korea, in 2012 and 2016, respectively. She is a currently assistant professor at the Department of Information

Security, Pai Chai University. From 2016 to 2018, she was a researcher at the Department of Computer Science and Engineering at the University at Buffalo, State University of New York, USA. She was a research professor at SCH Convergence Science Laboratory, Soonchunhyang University from 2019 to 2020. Her research interest includes computer security, privacy-preserving data mining and data science in general.