

# 캐릭터 이름을 이용한 MMORPG 봇 탐지 기법

강성욱\*, 이은조\*\*

## 요약

온라인 게임에서 불법 프로그램을 이용한 게임 봇을 대규모로 운영하는 전문 사설 업체를 속칭 ‘작업장(Gold Farming Group, GFG)’이라고 부른다. 기존에 작업장에서 운영하는 게임 봇은 24시간 쉬지 않고 반복적인 파밍을 통해 수익을 극대화하는 전략을 취했으나 최근 온라인 게임의 계정 가입이 쉬워지고 무료 플레이가 보편화되면서 개개의 게임 봇 계정이 수행하는 플레이 시간이나 취득 재화 수준을 낮추는 대신 수만 개의 계정을 번갈아 가며 운영하는 방식으로 변하고 있다. 이로 인해 플레이 활동 패턴에 기반한 기존의 탐지 모델들이 점차 무력화되고 있으며 진입 초기에 게임 봇을 빠르게 탐지하고 제재하는 방안이 점차 중요해지고 있다. 우리는 게임 봇을 조기에 탐지하기 위한 방안으로 계정 및 캐릭터의 이름이 갖는 특성을 활용한 게임 봇 탐지 기법을 제안한다. 제안한 기법의 유효성을 검증하기 위해 북미에서 서비스 중인 엔씨소프트의 MMORPG인 ‘블레이드 앤 소울’의 약 20만 개 계정 정보를 이용해 탐지 성능을 측정하였다. 실험에 의하면 캐릭터 이름에 대해 간단한 나이브 베이즈 분류기를 적용하는 것만으로도 AUC 기준으로 약 0.901의 성능을 기록하였다.

## I. 서론

온라인 게임의 인기와 상업적 성공의 이면에는 계정 도용, 사기, 사설 서버 등의 어두운 면도 공존하고 있다. 이중에서 소위 ‘오토’ 혹은 ‘게임 봇’이라고 부르는 불법 프로그램을 이용한 게임 플레이는 가장 심각한 보안 위협 중 하나이다. 특히 이런 ‘게임 봇’을 대규모로 운영하는 전문적인 사설 업체들이 존재한다. 이들은 속칭 ‘작업장’이라고 부르는데 과도한 상업 행위로 인해 게임 유저의 플레이를 방해할 뿐만 아니라 게임 회사에 상업적인 피해를 유발한다.

게임 회사에서는 이런 ‘작업장’을 단속하기 위해 다양한 방법을 활용하고 있다. 가장 직접적으로는 불법 프로그램을 탐지하는 모듈을 게임 클라이언트 프로그램에 탑재하거나 서버와 클라이언트가 주고 받는 통신 패킷의 위변조를 감시하는 방법이 있다. 여기서 더 나아가 게임 로그를 이용한 게임 봇 탐지 기법도 많이 활용되고 있는데 특히 최근에 기계 학습 기법이 널리 확산되고 발전하면서 게임 봇의 플레이 패턴을 학습하여 탐지하는 분류 모델이 많이 연구되고 있으며 실전에 적용하는 경우도 점차 확대되고 있다.

게임 로그의 정밀도가 점차 향상되고 기계 학습 기법

이 발전하면서 게임 봇 행위를 탐지하는 분류 모델의 정확도는 점차 향상되고 있다. 그럼에도 불구하고 게임 봇을 이용하는 행위가 쉽게 근절되지 못하고 있다. 그 이유는 먼저 게임 봇을 가장 많이 사용하는 작업장들이 탐지를 우회하기 위해 게임 회사 못지않게 적극적이고 다양한 방법을 동원하고 있기 때문이다. 또한 게임 봇 제재는 오탐으로 인한 비용이 매우 큰 문제가 있다. 즉, 게임 회사에서 일반 유저를 게임 봇으로 오인하고 잘못 제재할 경우 발생 되는 서비스 운영 비용 및 법적 분쟁으로 인한 비용은 게임 봇 탐지 및 제재를 보수적으로 수행하게 만드는 주요 요인이다.

한편, 최근 개인 정보 보호와 서비스 진입 장벽을 낮추기 위한 노력으로 인해 회원 가입이 점차 쉬워지면서 동일한 사람이 다수의 계정을 쉽게 생성할 수 있게 되었다. 또한 최근의 게임 트렌드는 게임 패키지를 구매하거나 월정액을 통해 서비스하는 대신 무료로 게임을 플레이할 수 있게 하는 대신 더 많은 서비스나 편의 기능을 활용하기 위해선 유료 아이템을 구매하도록 사업 정책이 바뀌고 있다. 이런 사업 모델을 ‘부분 유료화(free-to-play)’ 모델이라고 부른다. 대부분의 모바일 게임은 이런 사업 모델을 갖고 있으며 기존에 월정액 서비스를 제공하던 PC MMORPG 역시 점차 부분 유료

\* 엔씨소프트 (swkang@ncsoft.com)

\*\* 엔씨소프트 (gimmesilver@ncsoft.com)

화 모델로 전환하고 있다[1].

이런 변화는 작업장에도 영향을 주고 있다. 기존에 월정액 서비스를 이용하던 작업장들은 하나의 계정에서 최대한 이익을 얻기 위해 24시간 동안 거의 쉬지 않고 게임 봇들을 운영하며 게임 재화를 최대한 많이 획득하는데 집중하였다. 따라서 게임 봇 탐지를 위해 지속적으로 반복적인 플레이 패턴을 학습하여 탐지하는 기법이 많이 연구되었다[2-4].

그러나 온라인 게임의 사업 모델이 부분 유료화로 바뀌고 회원 가입이 쉬워지면서 게임 봇 계정을 생성하고 운영하는데 드는 비용이 매우 낮아지게 되었다. 이로 인해 작업장에서는 개개의 게임 봇 계정의 플레이 시간 및 재화 획득 수준을 낮추는 대신 수만 개의 계정을 번갈아 가며 운영하는 방식으로 변하고 있다. 보안상의 이유로 구체적인 수치를 밝힐 수는 없지만 이런 작업장들은 기존에 비해 게임 봇 계정 하나 당 하루에 획득하는 추정 수익이 1/10 이하로 줄어든 대신 한 번에 운영하는 게임 봇 규모는 몇 배 이상 늘어난 것으로 추정하고 있다.

이런 변화는 작업장을 단속하는 게임 회사에 다음과 같은 문제를 던진다.

첫째, 플레이 패턴을 이용한 기계 학습 모델을 효과적으로 만들기 어렵다. 분류 모델을 만들기 위해선 충분한 양의 데이터를 통해 패턴을 학습해야 하는데 게임 봇 계정 당 생성되는 플레이 데이터가 적기 때문에 그 특징을 충분히 학습하기 어렵다.

둘째, 짧은 플레이에서 패턴을 찾아 분류 모델을 적용하더라도 해당 모델을 통해 게임 봇을 탐지한 시점에는 작업장에서 이미 해당 계정을 더 이상 활용하지 않고 다른 계정을 이용하는 경우가 많이 발생하면서 한발 늦은 대응을 하게 된다.

셋째, 일반 유저 입장에서는 기존에 비해 더 많은 수의 게임 봇을 접하게 됨으로써 게임 회사에 대한 부정적인 이미지가 심화된다.

이런 문제를 해결하기 위해선 게임 봇에 대한 탐지 및 제재가 최대한 빠른 시점에 이루어질 수 있도록 탐지 기법과 제재 정책의 변화가 필요하다. 본 논문에서는 이를 위한 방안으로 캐릭터와 계정 이름을 이용한 탐지 기법을 시도해본 간단한 사례를 소개한다.

일반적으로 게임 사용자들은 자신이 플레이하는 캐릭터에 대한 감성적인 몰입도가 높기 때문에 자신의 개

성을 드러내는 이름을 짓는 경우가 많다. 따라서 많은 게임 회사에서는 게임 출시에 앞서 사전에 캐릭터 이름을 선점할 수 있는 이벤트를 제공하기도 한다[5]. 반면, 작업장의 경우 게임 봇에 대한 감성적인 몰입도가 매우 낮으며 대규모 계정을 생성해야 하는 한계로 인해 캐릭터 이름이나 계정 이름을 무작위 문자열로 생성하거나 혹은 관리 편의성을 위해 동일한 문자열 뒤에 일련의 번호를 붙이는 형태로 생성하는 경우가 대부분이다. 따라서 대다수의 게임 유저들은 캐릭터 이름만 보고도 게임 봇 여부를 쉽게 짐작하기도 한다. 따라서 기계 학습을 통해 일반 유저와 게임 봇 계정이 갖는 캐릭터 이름 패턴을 정확히 학습한다면 게임을 플레이하기 전에 이름만으로도 게임 봇을 탐지하여 게임 진입 전에 효과적으로 제재할 수 있을 것이다.

우리는 이를 위해 마르코프 체인 모델과 나이브 베이즈 분류기를 이용한 탐지 기법을 시도해 보았다. 기법의 유용성을 검증하기 위해 현재 북미에서 서비스 중인 엔씨소프트의 무협 MMORPG인 ‘블레이드 앤 소울’의 약 20만개 계정과 캐릭터 이름 데이터를 이용했다. 우리 실험에 의하면 캐릭터 및 계정 이름만으로도 비교적 높은 정확도로 게임 봇을 판별할 수 있는 것으로 나타났다.

본 논문은 다음과 같이 구성되어 있다. 우선 2장에서는 게임 봇 탐지 및 우리가 사용한 기계 학습 기법과 관련된 기존 연구 사례를 소개한다. 3장에서는 우리가 사용한 이름 기반 게임 봇 탐지 기법의 세부적인 알고리즘에 대해 소개한다. 4장에서는 3장에서 소개한 기법을 이용해 실제 게임 계정 및 캐릭터 이름으로 게임 봇을 분류한 실험 결과를 정리했으며 마지막 5장에서는 우리가 제안한 기법이 갖고 있는 한계점 및 후속 연구 방향에 대해 소개한다.

## II. 관련 연구

기계 학습을 이용한 게임 봇 탐지에 대한 연구는 2000년대 중반부터 본격적으로 시작했다.

Thawonmas 등[6]은 게임 봇이 일반 유저에 비해 같은 행동을 반복하는 경향이 높은 특징을 이용하여 게임 봇을 탐지하는 기법을 연구한 초기 논문이다. 이후 게임 봇의 반복성을 이용한 연구는 다양하게 이뤄졌는데 Mitterhofer 등[7], Chen 등[2], Kesteren 등[3]은 게임

봇의 이동 경로의 반복성을 이용한 탐지 기법을 제안한 사례이며 Lee 등[4]은 행동 패턴의 반복성을 자기 유사도라는 관점에서 연구한 사례이다. 위 연구들은 기존에 PC MMORPG 게임 봇 탐지에 적합한 기법인 반면, 앞서 언급했듯이 작업장의 게임 봇 활용 정책의 변화로 인해 그 효과가 점차 감소하고 있다. 특히 최근 모바일 게임의 경우 자동 플레이 기능을 기본적으로 탑재하고 있기 때문에 일반 유저와의 차별성이 사라지는 문제도 있다.

Ahmad 등[8]은 'EverQuest II' 라는 게임에서 게임 봇을 탐지하기 위해 다양한 분류 모델 알고리즘의 성능을 테스트하였다. 특히 이 논문에서는 게임 봇의 유형을 'gatherer', 'banker', 'dealer', 'marketer' 로 구분하고 각각의 특징을 분석하였다. 이런 게임 봇의 유형을 분류하고 각각의 특징에 맞게 탐지하려는 연구로는 Kang 등[9]과 Chung 등[10] 이 있다. [9]는 파티 활동에 특화된 게임 봇 유형에 대한 탐지 기법을 제안하고 있으며 [10]은 군집화 알고리즘을 이용해 게임 캐릭터들의 유형을 먼저 나눈 후 각 유형별 분류 모델을 별도로 만들어서 다양한 게임 봇 유형을 탐지하는 기법을 제안하고 있다. 이런 연구들은 플레이 유형에 따른 정교한 탐지가 가능하다는 점에서 참고할 만한 연구 사례들이다. 그러나 유형별 분류 모델을 적용하기 위해선 충분한 양의 플레이 기록이 필요하기 때문에 게임 봇이 게임을 시작하고 어느 정도 성장하기 전에는 탐지 및 학습이 어렵다는 한계가 있다.

작업장은 효율성을 높이기 위해 여러 개의 게임 봇이 분업화를 통해 협업하는 특징도 있다. 이런 특성을 갖는 게임 봇을 탐지하기 위해 네트워크 분석 기법을 적용한 연구들이 있다. Keegan 등[11], Kwon 등[12], Woo 등[13]이 이렇게 네트워크 분석을 이용한 작업장 탐지 기법을 연구한 대표적인 사례들이다. 네트워크 분석을 이용한 게임 봇 탐지 기법은 다수의 계정을 동시에 운용하는 작업장을 탐지하기에 적합한 방법이다. 특히 이렇게 집단 단위로 특징을 추적하고 탐지할 경우 이후 새로 생성되어 게임에 진입하는 게임 봇의 경우에도 기존에 탐지한 집단과의 관계를 추적함으로써 초기에 탐지가 가능하다. 그러나 이 기법은 기존 PC MMORPG처럼 유저 간의 사회 관계가 활발하고 관련 콘텐츠가 풍부한 게임에서는 효과적인 반면 최근 모바일 게임과 같이 캐릭터 사이에 상호 작용하는 콘텐츠가 비교적 적은

경우에는 효과가 떨어지는 한계점도 존재한다.

게임 봇과 일반 유저의 게임 플레이에 대한 동기가 다른 것에 주목하여 연구한 사례들도 있다. 이재혁 등[14]은 심리학에서 인간의 욕구에 대해 정리한 ERG 이론을 게임 봇 탐지에 활용한 사례이다. 게임 봇은 일반 유저와 달리 재화 획득을 통한 수익 창출 목적으로 플레이하기 때문에 이로 인한 동기배이션의 차이를 분류 모델의 특질로 사용한 방법이다. Kang 등[15]은 유저 간의 채팅 메시지 분석을 통해 게임 봇을 탐지하는 기법을 제안하였다. 이것 역시 일반 유저와 게임 봇이 갖고 있는 채팅의 목적이 다른 것에 초점을 맞춘 연구 사례이다. 위 연구들은 모두 게임 봇과 일반 유저의 게임 플레이 동기가 다르므로 인한 차이를 분류 모델의 특질로 이용한다는 점에서 우리 연구와 유사점이 있다.

마지막으로 본 연구에서 게임 봇을 분류할 때 사용한 알고리즘인 나이브 베이즈 분류기는 스팸 메일을 분류할 때 오래전부터 사용되던 알고리즘이다. 이와 관련해서는 McCallum 등[16], Chen 등[17], Kim 등[18]이 좋은 참고 사례이다.

### III. 탐지 알고리즘

#### 3.1. 마르코프 체인 모델

마르코프 체인은 어떤 상태가 이전 상태에 의해서만 결정된다고 가정하고 만든 상태 전이 확률을 이용해서 정의한 확률 모델을 말한다. 비록 현실을 많이 단순화한 모델이지만 비교적 적은 비용으로 어떤 초기 상태에서부터 시작된 상태 시퀀스의 확률을 계산할 수 있다. 만약 알파벳 a부터 h까지 8개의 상태가 존재할 때 'a'라는 상태에서 'b' 라는 상태로 전이하는 확률을  $P(b|a)$ 라고 하면 이것은 a라는 상태에서 다른 상태로 넘어가는 모든 사건 중 b라는 상태로 넘어가는 사건의 비율로 구할 수 있다. 이를 수식으로 표현하면 다음과 같다.

$$P(b|a) = \frac{\text{count}(a \rightarrow b)}{\text{count}(a \rightarrow *)}$$

위 식에서  $\text{count}(a \rightarrow *)$ 는 이전 상태가 a인 모든 사건의 횟수를 의미하며  $\text{count}(a \rightarrow b)$ 는 이전 상태가 a이고 이후 상태가 b인 사건의 횟수를 의미한다.

또한 초기 상태의 경우 이전 상태가 존재하지 않기 때문에 이 경우에는 전체 시퀀스 데이터에 대해서 초기 상태별 비율을 이용한다. 이를 수식으로 표현하면 다음과 같다.

$$P_{init}(a) = \frac{count_{init}(a)}{count(*)}$$

위 식에서  $count(*)$ 는 전체 시퀀스 데이터의 개수를 의미하며  $count_{init}(a)$ 는 각 시퀀스의 초기 상태가  $a$  인 데이터의 개수를 의미한다.

우리는 주어진 학습 데이터를 이용해서 어떤 문자 다음에 오는 문자의 횟수를 집계함으로써 상태 전이 확률 매트릭스를 만들 수 있다. 예를 들어 위에서 예시로 든 상태 전이 확률 매트릭스가 그림 1과 같다고 하자.

이 상태 전이 확률 매트릭스를 이용하여 ‘bahgffhaa’ 라는 단어가 등장할 확률은 아래와 같이 계산할 수 있다. 참고로 아래 예시에서는 알파벳 ‘b’의 초기 확률이 1/3 이라고 가정했다.

$$\begin{aligned} P(o_1 = b, o_2 = a, o_3 = h, o_4 = g, o_5 = f, o_6 = f, o_7 = h, o_8 = a, o_9 = a) \\ = P(b) \times P(a|b) \times P(h|a) \times P(g|h) \times P(f|g) \times P(f|h) \times P(h|f) \\ \times P(a|h) \times P(a|a) \\ = \frac{1}{3} \times \frac{2}{3} \times \frac{5}{17} \times \frac{2}{11} \times \frac{1}{2} \times \frac{8}{11} \times \frac{1}{11} \times \frac{4}{11} \times \frac{11}{17} \\ = 0.9243 \times 10^{-4} \end{aligned}$$

위에서 설명한 마르코프 체인 모델을 이용해서 게임 붓을 탐지하는 절차는 다음과 같다.

1) 기존에 확보한 일반 유저와 게임 붓의 이름 목록을 이용해서 각각에 대해 문자 상태 전이 확률 매트릭스를 생성한다.

2) 탐지 대상이 되는 유저 A의 계정명과 캐릭터 이름에 대해 앞서 생성한 상태 전이 확률을 이용해 발생 확률을 계산한다. 이 때 각 상태 전이 확률을 그대로 이용해서 곱셈 연산을 할 경우 시스템에서는 언더 플로우 현상이 발생할 수 있으므로 로그를 취한 후 덧셈 연산

$$A = \begin{pmatrix} 11/17 & 1/17 & 0 & 0 & 0 & 0 & 0 & 5/17 \\ 2/3 & 0 & 0 & 0 & 0 & 0 & 0 & 1/3 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2/11 & 8/11 & 0 & 1/11 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 4/11 & 0 & 0 & 0 & 0 & 1/11 & 2/11 & 4/11 \end{pmatrix}$$

(그림 1) 마르코프 체인 상태 전이 확률 매트릭스

을 수행한다. 여기서는 게임 붓 상태 전이 확률 매트릭스를 이용한 확률을  $P_b(A)$ 라고 하고 일반 유저 상태 전이 확률 매트릭스를 이용한 확률을  $P_h(A)$ 라고 한다.

3) 위에서 구한 두 값의 차 즉,  $P_b(A) - P_h(A)$ 를 구한다. 이 값이 클수록 게임붓일 확률이 높음을 의미한다.

### 3.2. 나이브 베이즈 분류기

나이브 베이즈 분류기는 각각의 특성이 서로 독립이라는 가정 하에 베이즈 정리를 이용하여 입력 데이터가 속할 카테고리의 확률을 계산하여 어떤 카테고리로 분류할지 결정하는 알고리즘이다. 어떤 입력 데이터 D에 대해서 해당 데이터가 C 클래스로 분류될 확률  $P(C|D)$ 는 베이즈 정리를 이용하면 아래와 같은 식으로 구할 수 있다.

$$P(C|D) = \frac{P(C) \times P(D|C)}{P(D)}$$

나이브 베이즈 분류기는 이러한 베이즈 정리를 이용해 분류하고자 하는 대상의 각 카테고리에 속할 확률을 측정한 후 그 확률이 가장 큰 쪽으로 분류하는 방법이다. 이를 식으로 표현하면 아래와 같다.

$$\begin{aligned} C_{MAP} = \operatorname{argmax}_{c \in C} P(cd) = \operatorname{argmax}_{c \in C} \frac{P(d|c) \times P(c)}{P(d)} \\ \approx \operatorname{argmax}_{c \in C} P(d|c) \times P(c) \end{aligned}$$

그런데 만약 학습 데이터에 없는 새로운 단어가 등장할 경우 위 계산에 사용되는 확률  $P(d|c)$ 는 무조건 0이 되는 문제가 있다. 따라서 이런 문제를 제거하기 위해  $P(d|c)$ 를 계산할 때 빈도에 +1을 해주는 Laplace smoothing 기법을 이용한다.

나이브 베이즈 분류기를 이용한 게임 붓 탐지 절차는 아래와 같다.

1) 각 캐릭터나 계정 이름을 n-gram을 이용해 분리한다. 본 논문에서는 n을 2부터 4까지 바꿔가며 성능을 측정한 후 가장 성능이 잘 나오는 n=4를 최종 모델로 적용하였다.

2) 각 n-gram에 대해서 빈도수를 구하여 확률을 계산한다. 이 때 위에서 언급한 Laplace smoothing을 적용한다.

3) 확률을 계산할 때 문자열이 길어질수록 확률이 과도하게 0에 가까워지는 문제를 방지하기 위해 확률을 계산할 때는 로그를 취한 후 덧셈 연산을 수행한다.

4) 게임 봇과 일반 유저의 학습 데이터를 이용해 각각 구한  $P_b$  와  $P_h$  의 차이값을 구한다. 이 차이가 클수록 게임 봇일 확률이 높음을 의미한다.

#### IV. 실험 및 평가

##### 4.1. 데이터 셋

우리가 실험에 사용한 데이터는 북미에서 현재 서비스 중인 ‘블레이드 앤 소울’이라는 온라인 게임의 계정명과 캐릭터명 목록이다. 2016년 4월부터 5월 사이에 게임에 접속한 유저들을 대상으로 했으며 이 기간에 게임 봇으로 탐지되어 제재된 계정과 그렇지 않은 계정 각각 101,149개씩을 추출하여 실험에 사용하였다. 이 데이터에서 다시 게임 봇과 정상 유저를 각각 91,149개씩 추출하여 학습 데이터로 구성하였고 나머지 각각 10,000개씩의 데이터를 테스트에 사용했다. 참고로 하나의 계정은 여러 개의 캐릭터를 보유할 수 있는데 우리는 이 중에서 가장 레벨이 높은 메인 캐릭터 정보만 사용했다 (표 1 참조).

캐릭터 이름에는 알파벳과 공백 문자가 허용되며 계정명에는 알파벳과 숫자 및 몇 가지 특수 기호가 허용된다. 우리는 상태의 종류를 단순화하기 위해 특수 기호는 모두 공백문자로 치환하여 사용했다. 따라서 캐릭터명은 총 27개(알파벳 26개 + 공백문자 1개), 계정명은 총 37개(알파벳 26개 + 숫자 10개 + 공백문자 1개)의 문자 상태가 존재한다.

[표 1] 실험 데이터 정리 표

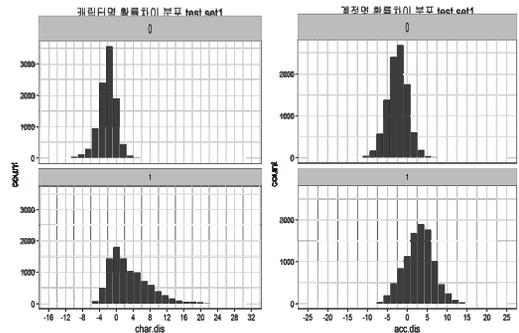
	Game bot	Human
Training set	91,149	91,149
Test set	10,000	10,000
Total	101,149	101,149

##### 4.2. 마르코프 체인 모델

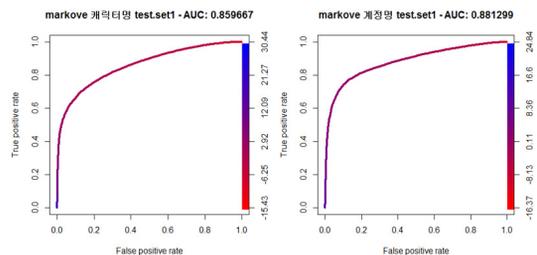
마르코프 체인 모델 생성을 위해 먼저 계정명과 캐릭터명 각각에 대해 전이 확률 테이블에 들어갈 상태를

정의 한다. 이 상태는 4.1.에서 언급했듯이 캐릭터명은 27개, 계정명은 37개의 상태로 이뤄진다. 상태 전이 확률 매트릭스는 게임 봇과 일반 유저 각각 생성하게 되며 따라서 총 4개의 매트릭스(게임 봇 캐릭터명, 게임 봇 계정명, 일반 유저 캐릭터명, 일반 유저 계정명)가 생성된다. 그림 2는 이렇게 생성한 상태 전이 확률 매트릭스를 이용해서 게임 봇과 일반 유저의 캐릭터명, 계정명의 확률을 구한 후 게임 봇 확률에서 일반 유저 확률을 뺀 차이값에 대한 분포를 비교한 결과이다. 그림에서 ‘0’이라고 되어 있는 것이 일반 유저이며 ‘1’이라고 되어 있는 것이 게임 봇이다. 차이값이 클수록 게임 봇일 확률이 높다는 것을 의미하는데 그림을 보면 게임 봇이 일반 유저에 비해 상대적으로 오른쪽에 분포가 쏠리는 것을 확인할 수 있다.

그림 3은 이 차이값을 이용하여 성능을 측정한 Receiver Operating Characteristics(ROC) 그래프이며 표 2는 이 ROC 그래프를 통해 구한 AUC를 비교한 자료이다. 캐릭터명에 비해 계정명을 이용한 분류 모델이 좀 더 높은 성능을 보여준다. 그 이유는 계정명이 캐릭터명에 비해 좀 더 많은 상태값을 갖고 있기 때문에 정교한 구분이 가능했던 것이라고 추정한다.



[그림 2] 게임 봇과 일반 유저 확률 분포 비교



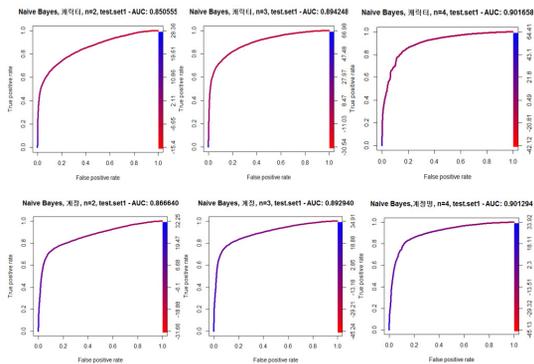
[그림 3] 마르코프 체인 모델의 ROC 그래프

[표 2] 마르코프 체인 모델의 AUC

	캐릭터명	계정명
AUC	0.859667	0.881299

### 4.3. 나이브 베이즈 분류기

나이브 베이즈 분류기에서는 문자열을 먼저 n-gram으로 전처리하여 이용한다. 우리는 n을 2부터 4까지 바꿔며 성능을 측정하였다 (그림 4와 표 3 참조). 각각의 ROC 그래프는 그림 4와 같다. 이 자료에 의하면 n=4일 때 가장 높은 성능을 보이며 이 때의 AUC는 캐릭터명과 계정명 모두 약 0.901이 나오는데 이는 마르코프 체인 모델에 비해 조금 더 우수한 결과이다.



[그림 4] 나이브 베이즈 분류기 ROC 그래프

[표 3] 나이브 베이즈 분류기의 AUC

	캐릭터명	계정명
AUC (n=2)	0.850555	0.866640
AUC (n=3)	0.894248	0.892940
AUC (n=4)	0.901658	0.901294

### 4.4. 실험 결과 정리

표 4는 위 분류 모델을 이용해 분류한 게임 붓과 일반 유저의 캐릭터명과 계정명 샘플이다. 일반 유저는 대부분 사람 이름이나 사전에 나오는 단어의 조합으로 구성된 반면 게임 붓은 의미를 알 수 없는 임의 문자의 조합으로 이뤄져 있는 것을 확인할 수 있다. 한 가지 재미있는 점은 캐릭터명의 경우 일본 이름이 많다는 점이다. 사용한 게임 데이터가 북미 서비스임에도 불구하고 이

[표 4] 게임 붓과 일반 유저의 이름 비교

계정명	
게임 붓	일반 유저
vcjtjpf__721862	rock__peace__
txolcv__739517	mistyfoster_and__
nbo__sr084062842	simon-__stein
hakbvqn__0257428	metalheart____
kbzbcyef__79428	alex.__toss
demxy__173951	__ziocalazans__
eqsfncj__914631	benjamin.____
k26479610__u040	andrea.____1994
ihldiwp7__84062	julian.____
suvfrjxcu__4062	lukinhas.____
캐릭터명	
fxskfx__objto	hokagekakashi__
xskcs__ogbtog	kawaii yuuki__
zrmd__vqidv	__yoshinatsu
xskfxskcx__	__shiraki
__tlgxskcspkc	miki hoshii__
__kfxskfxpkcp	yuki __shiro
tog__xskcxpkc	yoki aki__
__kfxskfxpkc	tsukuna __
__xpkfxobjtoj	suzuya ryun__
kfxr__jbowjb	sakura aki__

런 특성을 보이는 이유는 아마 ‘블레이드 앤 소울’의 게임 특성이나 작화 스타일 때문일 것이라 추정한다. 참고로 개인 정보 보호를 위해 각각의 문자열에서 일부를 마스킹 처리하였다.

한편, 실험에서 가장 높은 성능을 보인 것은 캐릭터 이름에 대해서 나이브 베이즈 분류기를 적용한 것이며 이 때의 AUC는 약 0.901이다. 기준에 동일 게임을 대상으로 게임 플레이 로그를 이용하여 만들었던 정교한 탐지 모델[4]의 AUC가 약 0.940 이었던 것과 비교해 보면 낮지 않은 수치이다. 특히 이번 실험의 목적은 컨셉 검증이었기 때문에 정교한 튜닝을 하지 않았음에도 불구하고 실전에 사용하고 있는 모델과 성능 차이가 많지 않다는 점은 인상적인 결과라 생각한다. 따라서 캐릭터나 계정의 이름을 이용하여 게임 붓을 탐지하는 기법이 충분히 실제 서비스에 적용 가능한 방법임을 입증하는 사례라고 생각한다.

## V. 결 론

온라인 게임에서 불법 프로그램을 이용한 게임 봇은 사용자와 게임 회사에게 직간접적으로 피해를 주기 때문에 건전한 서비스 운영을 위해 해결해야 할 중요한 보안 문제 중 하나이다. 특히 대규모 게임 봇을 전문적으로 운영하는 작업장은 게임 서비스의 존폐에 영향을 주는 중대한 보안 위협이기 때문에 게임 회사에서는 다양한 기법을 이용해 작업장을 제재하고 있다.

기존에 작업장은 게임 봇 계정 하나 당 최대한 오랜 시간 반복적인 파밍을 통해 재화를 취득하는 방식을 사용하였으나, 최근 게임 트렌드 변화로 인해 계정 생성이 쉬워지고 관리비가 거의 들지 않게 되면서 짧은 시간 동안 소규모 재화를 빠르게 취득하되 이런 계정들을 기존보다 훨씬 대규모로 운영하는 방식으로 변화하였다. 이로 인해 기존에 게임 플레이 패턴을 탐지하는 기계 학습 모델이 무력화되고 있다. 따라서 이런 작업장의 운영 정책에 대응하기 위해선 게임 진입 전 혹은 초기에 빠르게 게임 봇을 탐지하여 제재하는 방법이 필요하다.

본 논문에서는 마크로프 체인과 나이트 베이스 분류기를 이용해서 계정명 및 캐릭터명만으로 게임 봇을 탐지하는 기법을 제안하였다. 제안한 방법의 유효성을 검증하기 위해 북미에서 현재 서비스 중인 엔씨소프트의 MMORPG인 '블레이드 앤 소울'의 약 20만개 계정 정보를 이용했다. 빠른 컨셉 검증이 목적이었기 때문에 특별한 튜닝을 하지 않았음에도 불구하고 기존 탐지 모델에 비해 뒤지지 않는 정확도로 게임 봇을 분류할 수 있었다.

이 방법은 계정 및 캐릭터가 생성된 직후에 바로 활용이 가능하며 데이터 처리에 대한 비용이 매우 적기 때문에 게임 서비스 초기에 작업장에 대한 빠른 대응 목적으로 사용하기 적절할 것으로 생각한다. 또한 게임 콘텐츠 및 장르 특성에 대한 의존도가 거의 없기 때문에 다양한 게임에 활용 가능한 장점도 있다. 반면 캐릭터나 계정 이름만으로 게임 봇으로 단정하는 것은 제재 근거가 미약하며 탐지 우회가 비교적 쉽다는 단점도 있다. 따라서 이런 단점을 보완하기 위한 좀 더 정교한 분석 및 제재 정책을 고안하는 것이 필요할 것이다. 더 나아가 이 기법을 제재에 직접적으로 적용하기 보다는 다른 분류 모델과 함께 사용함으로써 단점을 보완하고 기존 기법의 성능을 개선할 수도 있을 것이다.

향후에는 이 기법을 다양한 게임에 테스트해 보는 한편 다른 분류 모델과의 조합을 통해 단점을 보완하고 성능을 개선하기 위한 방안을 연구할 예정이다.

## 참 고 문 헌

- [1] 남혁우, “블레이드앤소울, 이용권 폐지 부분 유료화 전환,” [http://www.zdnet.co.kr/news/news\\_view.asp?article\\_id=20161119223945](http://www.zdnet.co.kr/news/news_view.asp?article_id=20161119223945), Nov. 2016.
- [2] K.-T. Chen, A. Liao, H.-K. K. Pao, and H.-H. Chu, “Game Bot Detection Based on Avatar Trajectory,” in *Entertainment Computing ICEC 2008*, vol. 5309, pp. 94-105, 2009.
- [3] M. van Kesteren, J. Langevoort, and F. Grootjen, “A Step in the Right Detecting: Bot Detection in MMORPG using Movement Analysis,” in *The 21<sup>st</sup> Benelux Conference on Artificial Intelligence*, 2009.
- [4] Lee, E., Woo, J., Kim, A., Mohaisen, A., and Kim, H., K., “You are a game bot!: uncovering game bots in MMORPGs via self-similarity in the wild,” in *NDSS 2016*, Feb. 2016.
- [5] 김범수, “리니지M, 캐릭터 사진 생성 8일만에 마감...5월 26일 서버 20개 추가,” [http://biz.chosun.com/site/data/html\\_dir/2017/05/25/2017052501507.html](http://biz.chosun.com/site/data/html_dir/2017/05/25/2017052501507.html), May 2017
- [6] R. Thawonmas, Y. Kashifuji, and K.-T. Chen, “Detection of MMORPG Bots based on Behavior Analysis,” in *Advances in Computer Entertainment Technology Conference*, pp. 91-94, Dec. 2008.
- [7] S. Mitterhofer, C. Kruegel, E. Kirda, and C. Platzer, “Server-side Bot Detection in Massively Multiplayer Online Games,” *Security and Privacy, IEEE*, vol. 7, no. 3, pp. 29-36, May 2009
- [8] M. A. Ahmad, B. Keegan, J. Strivastava, D. Williams, and N. Contractor, “Mining for Gold Farmers: Automatic Detection of Deviant Players in MMOGs,” in *Computational Science and Engineering International Conference*, vol. 4, pp. 340-345, Aug. 2009.

- [9] A. R. Kang, J. Woo, J. Park, and H. K. Kim, "Online game bot detection based on party-play log analysis," *Computers & Mathematics with Applications*, vol. 65, no. 9, pp. 1384-1395, 2013.
- [10] Y. Chung, C. yong. Park, N. ri, Kim, H. Cho, T. Yoon, H. Lee, and J. -H. Lee, "Game Bot Detection Approach Based on Behavior Analysis and Consideration of Various Play Style," in *Journal of ETRI*, vol. 35, no. 6, pp. 1058-1067, Dec. 2013.
- [11] B. Keegan, M. A. Ahmad, J. Srivastava, D. Williams, and N. Contractor, "Dark Gold: Statistical Properties of Clandestine Networks in Massively Multiplayer Online Games," in *Social Computing(SocialCom)*, IEEE Second International Conference, pp. 201-208, Aug. 2010.
- [12] H. Kwon, A. Mohaisen, J. Woo, Y. Kim, and H. K. Kim, "Crime Scene Reconstruction: Online Gold Farming Network Analysis," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 3, pp. 544-556, Nov. 2016.
- [13] J. Woo, A. Kang, and H. K. Kim, "The Contagion of Malicious Behaviors in Online Game," *Computer Communication Review*, vol. 43, pp. 543-544, 2013.
- [14] 이재혁, 강성욱, 김휘강, "모티베이션 이론을 이용한 온라인 게임 내 부정행위 탐지," *한국게임학회 논문지*, vol. 15, no. 4, pp. 69-78, Aug. 2015.
- [15] A. R. Kang, H. K. Kim, and J. Woo, "Chatting Pattern Based Game Bot Detection: Do They Talk Like Us," in *KSII Transactions on Internet and Information Systems*, vol. 6, no. 11, pp. 2866-2879, Nov. 2012.
- [16] A. McCallum, and K. Nigam, "A Comparison of Event Models for Naive Bayes Text Classification," in *AAAI-98 workshop on Learning for Text Categorization*, vol. 752, pp. 41-48, Jul. 1998.
- [17] J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature Selection for Text Classification with Naive

Bayes," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5432-5435, Apr. 2009.

- [18] S.-B. Kim, K. S. Hand, H. C. Rim, and S. H. Myaeng, "Some Effective Techniques for Naive Bayes Text Classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 11, pp. 1457-1466, Sep. 2006.

## 〈저자 소개〉



**강성욱 (Kang, Sung Wook)**  
정회원

2014년 2월 : 금오공과대학교 소프트웨어공학부 학사

2016년 2월 : 고려대학교 정보보호대학원 석사

2015년 12월~현재 : 엔씨소프트 데이터분석모델링팀

관심분야 : 데이터마이닝



**이은조 (Lee, Eun Jo)**

정회원

2002년 2월 : 숭실대학교 컴퓨터 학부 학사

2007년 2월 : 숭실대학교 정보통신대학원 석사

2015년 2월~현재 : 고려대학교 정보보호대학원 박사과정

2007년 5월~현재 : 엔씨소프트 데이터분석모델링팀 팀장  
관심분야 : 온라인 게임 보안, 데이터 마이닝