

## MMORPG 사용자 유형 분류를 통한 이탈 예측 모델 생성 및 평가

Constructing and Evaluating a Churn Prediction Model using Classification of User Types in MMORPG

---

저자 (Authors)	오세준, 이은조, 우지영, 김휘강 Sejoon Oh, Eunjo Lee, Jiyoung Woo, Huy Kang Kim
출처 (Source)	<a href="#">정보과학회 컴퓨팅의 실제 논문지 24(5)</a> , 2018.5, 220-226(7 pages) <a href="#">KIISE Transactions on Computing Practices 24(5)</a> , 2018.5, 220-226(7 pages)
발행처 (Publisher)	<a href="#">한국정보과학회</a> The Korean Institute of Information Scientists and Engineers
URL	<a href="http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07434425">http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07434425</a>
APA Style	오세준, 이은조, 우지영, 김휘강 (2018). MMORPG 사용자 유형 분류를 통한 이탈 예측 모델 생성 및 평가. 정보과학회 컴퓨팅의 실제 논문지, 24(5), 220-226
이용정보 (Accessed)	한신대학교 218.155.***.152 2019/07/19 17:00 (KST)

---

### 저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

### Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

# MMORPG 사용자 유형 분류를 통한 이탈 예측 모델 생성 및 평가

## (Constructing and Evaluating a Churn Prediction Model using Classification of User Types in MMORPG)

오 세 준 <sup>†</sup>      이 은 조 <sup>††</sup>      우 지 영 <sup>†††</sup>      김 휘 강 <sup>††††</sup>  
(Sejoon Oh)      (Eunjo Lee)      (Jiyoung Woo)      (Huy Kang Kim)

**요 약** 대규모 다중 사용자 온라인 롤플레이팅 게임(MMORPG)은 세계적으로 많은 사용자들이 즐기는 게임의 장르로 사용자들이 즐길 수 있는 다양한 콘텐츠를 제공해준다. 하지만 다양한 콘텐츠를 제공함에도 불구하고 일부 사용자들은 게임에서 이탈한다. 지속적으로 게임을 즐기는 사용자는 게임회사의 수익과 밀접한 관계가 있기 때문에, 사용자 이탈 예측은 매우 중요한 문제이다. 본 연구에서는 게임 사용자 유형을 기반으로 이탈을 예측하는 방법을 제안한다. 실제 MMORPG 데이터를 이용하여 ‘이탈자’를 정의하고, 서로 다른 특징을 보이는 사용자들을 군집화 알고리즘을 통해 5가지 유형으로 나누어 분류하였다. 실험 결과로 약 98.3%의 이탈자가 라이트 사용자형에 속하는 것을 확인하였다. 분류한 유형을 이탈 예측 모델 생성의 피쳐로 사용하고, 기계학습 알고리즘을 이용하여 이탈 예측 모델을 구축하였다. 이탈 예측 모델은 최대 85.7%의 accuracy와 72.3%의 F-measure의 성능을 보였다.

**키워드:** 사용자 이탈 예측, 사용자 유형 기반 분류, 기계학습 알고리즘, 이탈 예측 모델 평가

**Abstract** Massively Multiplayer Online Role Playing Games (MMORPG) have been the most enjoyed genre of games by users all over the world, because of the many enjoyable contents they provide. However, there still are some users leaving MMORPG, in spite of the various contents. Because the users, who are constantly engaged in a game, are directly connected to the profit of the game company, it is necessary to study churn prediction. In this study, we propose a churn prediction method based on the user types. Using actual MMORPG data, we define “Churn Users” and categorize users into 5 types by means of the clustering algorithm. The experiment shows that about 98.3% of the churn users are the “Light user type”. We then use user types as a feature and construct a churn prediction model with various machine learning algorithms. The churn prediction model shows a maximum accuracy of 85.7% and an F-measure of 72.3%.

**Keywords:** churn prediction, user type classification, machine learning algorithm, evaluating churn prediction model

<sup>†</sup> 학생회원 : 고려대학교 정보보호대학원  
sejoon222@korea.ac.kr

<sup>††</sup> 비 회 원 : 엔씨소프트  
gimmesilver@ncsoft.com

<sup>†††</sup> 정 회 원 : 순천향대학교 빅데이터공학과 교수  
jywoo@sch.ac.kr

<sup>††††</sup> 종신회원 : 고려대학교 정보보호대학원 교수(Korea Univ.)  
cenda@korea.ac.kr  
(Corresponding author인)

논문접수 : 2017년 10월 25일

(Received 25 October 2017)

심사완료 : 2018년 3월 15일

(Accepted 15 March 2018)

Copyright©2018 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.  
정보과학회 컴퓨팅의 실제 논문지 제24권 제5호(2018. 5)

## 1. 서론

대규모 다중 사용자 온라인 롤플레이 게임(MMORPG)은 세계적으로 많은 사용자들이 즐기는 게임 장르이다. 사용자는 자신의 캐릭터를 직접 생성하여 게임 디자이너가 설계한 시나리오를 따라가며 도전과제를 해결하고, 사냥을 하고, 아이템 강화 등을 통해 캐릭터를 육성한다. 그 과정에서 아이템 채집, 제작 등의 혼자서 즐길 수 있는 활동과 사용자 간의 전투(PvP), 파티 플레이, 아이템 교환 등의 사회적 활동을 즐긴다. 이렇게 MMORPG에는 다양한 콘텐츠들이 존재하고, 사용자들은 콘텐츠를 즐기며 여러 가지 행동을 하게 된다.

한편 일부 사용자는 시간이 지날수록 게임에 흥미를 잃어 지속적으로 게임을 즐기지 못하고, 결국 이탈하게 된다. 사용자의 이탈은 직접적으로는 게임 회사 수익의 감소로 이어지며, 간접적으로는 이탈한 사용자에게 의한 부정적인 입소문이 잠재적인 사용자와 기존 사용자들에게 영향을 줄 수 있다[1]. 따라서 온라인 게임 산업의 규모가 점점 커지고 체계화되면서 회사의 지속적인 성장과 수익 유지를 위해 게임 회사에서는 사용자 이탈을 예측하고 방지하기 위한 연구의 필요성이 점차 강조되고 있다[2].

MMORPG에서는 게임 내에서 수행한 사용자의 다양한 활동이 모두 로그로 기록된다. 이탈자의 게임 이탈 원인은 이탈자의 성향에 따라 조금씩 다를 것이며 그런 성향은 게임 내에서 어떤 활동을 주로 했는지 게임 활동 로그를 통해 드러날 것이다. 따라서 게임 사용자들이 게임 내에서 보여주는 활동을 기반으로 행동 유형을 먼저 분류한 후 이 정보를 활용한다면 주로 어떤 유형의 사용자들이 게임에 흥미를 쉽게 잃어버리고 이탈하는지를 좀 더 쉽게 찾을 수 있을 것이다. 더 나아가 이탈자들이 주로 많이 분포하는 유형이 있다면 해당 유형의 특징을 분석하고 이를 통해 이탈을 방지하기 위한 개선안을 찾아 게임에 반영함으로써 장기적으로 사용자의 기대 수명을 높이는데 활용할 수 있다.

본 연구에서는 엔씨소프트에서 개발하고 서비스하고 있는 MMORPG 중 하나인 아이온의 실제 데이터 셋을 이용하였다. 사용자의 몇 가지 주요 행동 로그를 피쳐로 이용하여 군집화 알고리즘을 적용하였고 이를 통해 총

5가지 유형으로 사용자를 분류하였다. 이렇게 분류한 사용자 유형 정보를 통해 이탈자가 많은 비중을 차지하는 유형은 무엇인지 확인하였으며 더 나아가 이 정보를 활용한 이탈 예측 모델을 생성하고 평가하였다.

본 연구는 2장에서 사용자 이탈에 관한 관련 연구와 그 방법들, 그리고 사용자 행동 분석에 관한 연구에 대해 기술하였다. 3장에서는 먼저 이탈 기준을 정의하고 본 연구에서 사용한 사용자 행동 유형 분류 및 이탈 예측 방법을 기술한다. 연구에서 사용한 방법론의 개략적인 흐름은 그림 1과 같다. 4장에서는 아이온의 로그 데이터를 이용한 사용자 유형 분류 및 이탈 예측 실험 결과를 소개하며 예측 모델에 대해서 True Positive Rate (TPR), False Positive Rate (FPR), Accuracy, Precision, Recall, F-measure를 이용해 그 성능을 평가하였다. 5장에서는 본 연구의 결론과 향후 연구 방향에 관하여 기술하였다.

## 2. 관련 연구

본 연구와 관련한 기존 연구를 표 1과 같이 크게 사용자 이탈 예측에 관한 연구와, 사용자 행동 유형 분석에 관한 연구로 분류하여 살펴보았다.

### 2.1 사용자 이탈 예측에 관한 연구

사용자 이탈 예측은 온라인 게임 분야에서 중요한 문제로 다루어져 왔다. Borbora와 Srivastava[3]는 이탈자와 지속적으로 게임을 즐기는 사용자를 비교하는 방법으로 사용자의 생활주기를 비교, 분석하는 방법으로 wClusterDist라고 부르는 거리 기반 분류 체계를 제안하였고, Hadji 등[4]은 무료로 플레이할 수 있는 게임에 대한 사용자 이탈 예측을 연구하였다. Borbora 등[5]은 사용자의 게임에 대한 동기에 성취감, 사회 활동, 몰입 측면에서 이론 중심의 접근과 데이터 중심의 접근 두 방법에 관하여 연구하였고, Hadden 등[6]은 신경망, 회귀 그리고 의사 결정 트리를 이용하여 사용자의 불평을 기반으로 이탈 연구를 수행하였다. Kawale 등[1]은 게임 내의 사회 활동과 사용자 이탈간의 관계를 분석하여 사용자 이탈이 사회 활동에 많은 영향을 받는 것을 발견하였다. Shores 등[7]은 다중 사용자 온라인 배틀 아레나(MOBA) 장르의 게임인 리그 오브 레전드에서 이탈자가 게임에 미치는 영향에 관한 연구를 통해 친구



그림 1 사용자 유형 기반의 이탈자 예측 모델

Fig. 1 Churn prediction model based on user type

표 1 관련 연구 분류

Table 1 Classification of related work

Category		Description
Churn prediction	Online game	- Life-cycle analysis using wClusterDist [3] - Free to play game analysis [4] - Two different models using clustering [5] - Complaints data analysis using decision tree [6] - Influence of churn prediction in MOBA game [7] - Comparison of machine learning techniques [8]
	Banking	- Comparison of machine learning techniques [8]
	Insurance	
	Telecommunication	- Social relationship analysis [9] - Churn prediction using boosting algorithm [10]
User behavioral pattern analysis	Bot detection	- Sequence analysis [11,12] - Classification of game character's growth type [13]
	Fraud detection	- Classification of identity theft case [14]

서비스와 같은 사회 활동적인 측면에서 게임을 발전시켜야 한다는 결론을 얻었다.

사용자 이탈은 온라인 게임 분야에서 뿐만 아니라 은행, 보험, 통신 등 다양한 분야에서 중요하게 생각되어 왔다[1,8]. Vafeiadis 등[8]은 앞서 언급한 여러 분야의 사용자 이탈 예측 문제를 위해 다양한 기계 학습 알고리즘의 성능을 비교했다. 특히 통신 분야에서 Dasgupta 등[9]은 사회적 관계가 이탈에 중요한 영향을 미친다는 것을 제시하였고, Lu 등[10]은 이탈 예측 분석을 위해 기계 학습 알고리즘 중 하나인 부스팅 알고리즘 사용을 제안하고 성능을 평가하였다.

## 2.2 사용자 행동 유형 분석에 관한 연구

Lee 등[11,12]은 게임 내 사용자 행동 로그 데이터를 기반으로 시퀀스를 추출하고 분석하는 방법에 대해 연구하여 게임 봇 탐지에 적용하였다.

한편 Lee 등[13]과 Choi 등[14]은 각각 사용자의 행동 유형을 기반으로 게임 봇과 계정 도용 행위를 탐지하는 방법에 대해 연구를 제안하였다.

## 3. 방법론

이 장에서는 이탈자 분류를 위해 사용한 세부 방법론을 기술한다. 절차는 크게 1) 이탈 기준 정의, 2) 사용자 행동 기반 유형 분류, 3) 이탈자 분류 모델링 및 평가로 이뤄진다.

### 3.1 이탈 기준 정의

정확한 이탈자 분류를 위해서는 합리적인 이탈 기준을 정해야 한다. 만약 이탈 기준을 지나치게 엄격하게 정하면 실제 이탈한 사용자를 비이탈자로 취급하게 되는 2종 오류를 높이 수 있고, 반대로 기준을 완화할 경우 이탈하지 않은 사용자를 이탈자로 판단하는 1종 오류를 높일 수 있다. 따라서 적절한 기준을 정하는 것은 실용적이고 정확한 이탈자 분류 모델을 만드는데 필수적이다[2].

이 연구에서는 Yang 등[15]이 사용했던 방법을 참고하여 'N일 동안 접속이 없는 사용자'를 이탈자로 정의한다. 이 때 적절한 N값을 찾기 위해 사용자의 미접속 기간에 따라 완전히 이탈하는 사용자의 비율을 측정 한 후 비율 변화가 둔화되는 변곡 시점을 N값으로 정하였다. 구체적인 수치 및 비율 변화에 대한 실험 결과는 4.1에서 소개한다.

### 3.2 사용자 행동 기반 유형 분류

사용자가 게임 내에서 수행한 활동을 기반으로 유형을 분류하기 위해 아이온에서 제공하는 다양한 콘텐츠를 사용하였다. 아이온에서 제공하는 콘텐츠의 종류는 다양하지만 가급적 범용성을 확보하기 위해 타 MMORPG에서도 일반적으로 제공하는 콘텐츠인 사냥, 던전 입장 횟수, 아이템 채집 및 제작, PvP 전투, 퀘스트 수행 횟수, 파티 활동 횟수를 피쳐로 선정하였다. 표 2는 사용한 피쳐 목록 및 그 의미를 정리한 표이다.

사용자 유형을 분류하기 위해 K-평균 군집화 알고리즘을 사용하였다. K-평균 군집화 알고리즘은 분석가가 분류할 유형 개수를 지정하면 해당 개수만큼의 중심점을 생성한 후 이 중심점을 기준으로 각 유형 내 분산을 최소화하는 방식으로 유형을 분류하는 알고리즘이다. 비록 분석가가 적절한 유형 개수를 직접 지정해줘야 하고 아웃라이어에 민감하다는 단점이 있지만 상대적으로 다른 알고리즘에 비해 대량의 데이터에 대해서도 빠른 성능을 갖고 있으면서도 평균적인 성능을 보장하는 장점을 갖고 있다.

### 3.3 이탈자 분류 모델링 및 평가

이탈자 분류 모델 생성에 사용한 피쳐는 표 3과 같다. 앞서 사용한 MMORPG에서의 일반적인 행위 이외에, 사용자의 행동의 빈도로 활동성을 나타내는 전체 로그 횟수(Log count)와 사용자 유형 분류를 추가적인 피쳐로 선정하였다.

표 2 군집화를 위한 피처 선정  
Table 2 Feature selection for clustering

Feature	Description
Hunting	# of hunting actions
Instance dungeon	# of instance dungeon play actions
Item gathering	# of item gathering actions
Item making	# of item making actions
PvP combat	# of player vs. player combat actions
Quest	# of quest actions
Party	# of party play actions

표 3 이탈자 분석을 위한 피처 선정  
Table 3 Feature selection for churn analysis

Feature	Description
Hunting	# of hunting actions
Instance dungeon	# of instance dungeon play actions
Item gathering	# of item gathering actions
Item making	# of item making actions
PvP combat	# of player vs. player combat actions
Quest	# of quest actions
Party	# of party play actions
Log count	# of all actions
User type	Classified user type

이탈자 분류 모델을 평가하기 위해서 다양한 기계학습 알고리즘을 이용하였다. 본 연구에서 실험한 알고리즘은 먼저 트리 기반 알고리즘 중 C4.5, 로지스틱 회귀와 의사 결정 트리 학습을 결합한 로지스틱 모델 트리(LMT), 다수의 결정 트리를 임의적으로 학습하는 방식의 앙상블 방법인 랜덤 포레스트, k개의 임의로 선택된 속성을 고려하는 트리를 학습하는 방식인 랜덤 트리가 있으며, 자료 분석을 위한 대표적인 분류 알고리즘인 서포트 벡터 머신(SVM)과 의사 결정 트리 알고리즘에서 효과적인 기술인 축소 오류 제거(REP)와 연관된 규칙을 기반으로 하는 알고리즘인 JRip도 사용했다. 각 알고리즘을 이용해 생성한 모델의 성능을 평가할 때는 10-fold cross validation을 이용하였으며, 이탈자에 대한 TPR, FPR, Accuracy, Precision, Recall, F-measure를 측정하였다. 각 성능 지표의 측정 방법은 식 (1)과 같다.

$$\begin{aligned}
 Precision &= \frac{TP}{TP+FP}, \quad Recall = TPR = \frac{TP}{TP+FN}, \\
 FPR &= \frac{FP}{FP+TN}, \\
 TP: & \text{ True Positive}, \quad FP: \text{ False Positive}, \\
 TN: & \text{ True Negative}, \quad FN: \text{ False Negative}, \\
 F\text{-measure} &= 2 \times \frac{precision \times recall}{precision + recall}
 \end{aligned}
 \tag{1}$$

## 4. 실험

### 4.1 이탈 기준 정의

이 연구에서는 한국의 유명한 온라인 게임 회사 중 하나인 엔씨소프트로부터 제공받은 아이온의 약 4억 6천만 건의 실제 로그 데이터를 사용하였다. 이 데이터는 2015년 12월 30일부터 2016년 4월 20일까지 16주 동안 약 4만 6천명의 실제 사용자 게임 활동 정보를 담고 있다.

먼저 이탈 기준을 정하기 위해 총 16주 동안 게임 활동을 한 사용자의 접속 기록을 이용하여 미접속 기간에 따른 실제 이탈자의 비율을 그림 2와 같이 누적 그래프로 살펴보았다. 이에 의하면, 이탈 기준을 길게 설정할수록 거의 일정한 비율로 이탈자의 비율이 증가하는데, 이탈 기준을 9주로 설정했을 때 이탈자의 비율이 10% 이상 급격히 증가하고, 이후 비율 변화가 둔화하였다. 따라서 본 연구에서는 마지막으로 접속한 날로부터 63일 동안 접속하지 않은 사용자를 이탈자로 정의하였다.

한편 활동 기간에 따라 사용자의 특성이 다를 수 있기 때문에 본 연구에서는 전체 16주 데이터 중 가장 이탈자가 많은 주의 데이터를 선정하는 작업을 수행하였다. 이 때 이탈 기준을 9주 이상 미접속으로 정했기 때문에 7주차 이후에 접속한 사용자는 분석 작업에서 제외했다. 그림 3은 제공받은 데이터 중 1주차부터 7주차까지 N주차에 이탈하는 이탈자의 비율을 N값을 변화시

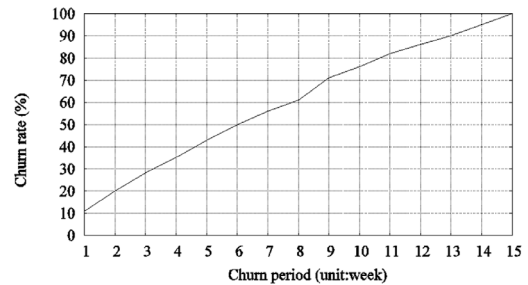


그림 2 이탈 기간별 이탈자의 비율  
Fig. 2 Churn rate by churn period

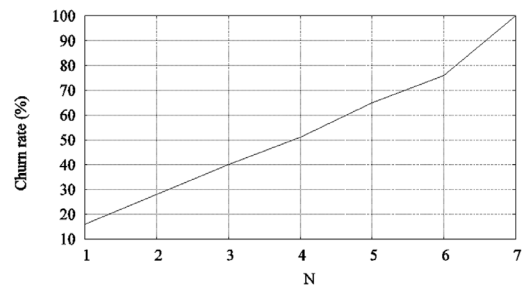


그림 3 데이터 기간별 이탈자의 비율  
Fig. 3 Churn rate by data period

표 4 사용자 행동의 군집화 결과  
Table 4 Clustering results of user action

Feature	Clustering results				
	1	2	3	4	5
Hunting	60	1,116	2,505	412	<b>4,295</b>
Instance dungeon	2	21	44	11	<b>57</b>
Item gathering	1	11	36	<b>93</b>	0
Item making	9	66	114	<b>3,037</b>	0
PvP combat	1	<b>12</b>	11	3	0
Quest	45	680	<b>1,784</b>	367	22
Party	11	99	161	38	<b>1,151</b>

키며 확인한 누적 그래프이다. 그림 3에서 보는 것과 같이 2016년 2월 10일부터 2016년 2월 17일까지의 제공받은 7주차 데이터에서 정의한 이탈자 모두가 이탈한 것과, 6주차 데이터에 비해 이탈자가 약 24% 증가한 것을 확인하였다. 따라서 7주차 기간 동안 활동한 10,260명의 사용자의 약 972만개의 로그 데이터를 실험에 사용하였다.

#### 4.2 사용자 유형 군집화

표 4는 K-평균 군집화 알고리즘을 사용하여 군집화한 각 군집의 중심점 데이터를 정리한 표이다. 이 결과를 통해 사용자를 표 5와 같이 1번 유형은 라이트 사용자형, 2번, 3번 유형은 다양한 콘텐츠형, 4번 유형은 아이템 제작형, 5번 유형은 사냥 위주의 파티 활동형으로 정의했다. 이 중 2번, 3번 유형의 다양한 콘텐츠형 사용자들은 활동 빈도에 따라 세부적으로 라이트, 헤비 사용자로 분류하였다.

먼저 1번 유형(라이트 사용자형) 사용자들은 다른 유형의 사용자들에 비해 모든 활동에서 전반적으로 낮은 수치를 보였다. 이들은 게임 활동을 거의 하지 않는 유형이라고 볼 수 있다. 다양한 콘텐츠형 사용자들은 다른 활동에 비해 사냥 활동과 퀘스트 수행 횟수가 많은 사용자 유형이다. 특히 2번 유형(다양한 콘텐츠형-라이트)

사용자들은 3번 유형(다양한 콘텐츠형-헤비) 사용자들과 비교했을 때 모든 활동에서 적은 수치를 기록하였으나 PvP 전투 활동에서 높은 수치를 기록하였다. 3번 유형(다양한 콘텐츠형-헤비) 사용자들은 모든 유형의 사용자들 중 퀘스트 수행을 가장 많이 하였고, 다른 모든 활동에서도 준수한 수치를 보였다. 이들은 다양한 콘텐츠를 즐기면서도 특히 퀘스트 완료를 목적으로 전반적으로 다양한 콘텐츠를 즐기는 유형이다. 4번 유형(아이템 제작형) 사용자들은 다른 유형의 사용자들에 비해 아이템 채집과 아이템 제작에서 많은 활동을 하였다. 이들은 아이템 채집, 사냥, 퀘스트 등의 활동을 통해 재료를 모으고 아이템 제작을 즐겨하는 유형이다. 5번 유형(사냥 위주의 파티 활동형) 사용자들은 다른 유형에 비해 파티 활동과 사냥 활동의 수치가 월등히 높았다. 아이템 제작, 채집, PvP 전투 활동을 거의 하지 않는 이들은 오로지 파티 사냥만을 즐기는 유형이다.

사용자 유형별 이탈자의 비율은 표 6에서 보는 것과 같다. 5번 유형 사용자는 총 612명 중 단 한명도 이탈하지 않았고, 2번 유형 사용자는 1,438명 중 29명이, 3번 유형 사용자는 472명 중 4명이 이탈하였다. 반면 1번 유형과 4번 유형 사용자의 이탈률은 각각 35.8%, 23.7%를 기록하여 전체 이탈자 2,793명 중 2,760명(전체 이탈자 중 98.8%를 차지)이 두 유형에 속하였다. 이것으로 보아 사용자의 행동 유형과 게임 이탈은 높은 상관성을 갖고 있다고 추정할 수 있다.

#### 4.3 사용자 이탈 분석

앞서 정의한대로 주어진 사용자 데이터에서 이탈자를 추출하여 총 10,260명의 사용자 중 2,793명을 이탈자로 정의하였고, 9개의 피처(사냥, 인스턴스 던전, 아이템 채집,

표 5 사용자 유형  
Table 5 User types

Clustering type	Description of the user type	
Type 1	Light user type	
Type 2	Various contents type	Light
Type 3		Heavy
Type 4	Item making type	
Type 5	Hunting with party play type	

표 6 사용자 유형별 이탈률  
Table 6 Churn rate by user type

User type	1	2	3	4	5
# of churn	2,746	29	4	14	0
# of users	7,679	1,438	472	59	612
Churn rate	35.8%	2.0%	0.8%	23.7%	0.0%

표 7 분류 알고리즘을 이용한 이탈 예측

Table 7 Churn prediction using different classification algorithms

Method	TPR	FPR	Accuracy	Precision	Recall	F-measure
C4.5	68.5%	7.8%	85.7%	76.6%	68.5%	72.3%
LMT	66.5%	7.2%	85.7%	77.6%	66.5%	71.6%
Random forest	70.9%	9.8%	85.0%	73.0%	70.9%	72.0%
Random tree	73.4%	14.0%	82.5%	66.2%	73.4%	69.6%
SVM	65.8%	6.8%	85.7%	78.3%	65.8%	71.5%
JRip	68.5%	8.1%	85.5%	76.0%	68.5%	72.0%

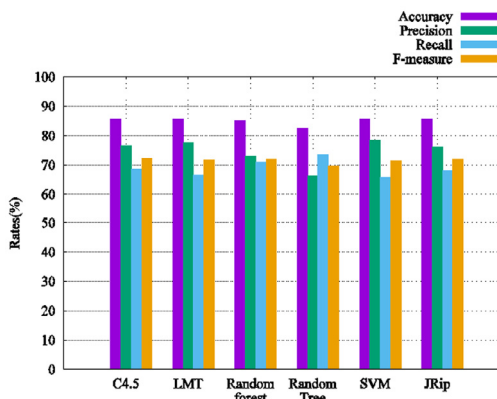


그림 4 분류 알고리즘 성능

Fig. 4 Classifier performance

아이템 제작, PvP 전투, 퀘스트, 파티, 총 로그의 수, 분류한 사용자 유형)를 기준으로 이탈 예측 모델을 생성하였다.

그림 4와 표 7은 알고리즘별 성능을 정리한 결과이다. 이 결과에 의하면 SVM 알고리즘이 가장 낮은 1종 오류를 갖는 반면 랜덤 트리 모델이 가장 낮은 2종 오류를 갖는다. 종합적으로 볼 때 accuracy나 F-measure가 가장 높은 의사 결정 트리(C4.5)가 가장 우수한 성능을 갖는다고 볼 수 있다.

## 5. 결론

본 연구에서는 현재 서비스 중인 아이온의 로그 데이터를 이용하여 이탈 기준을 정의하고 사용자들의 행동 유형을 5가지의 유형으로 분류하여 유형별 이탈자의 비율을 비교하였다. 실험에 의하면 사용자 유형에 따라 이탈률이 크게 차이가 나며 따라서 사용자의 게임 내 행동 유형이 이탈과 높은 상관관계를 갖는다고 추정할 수 있었다. 또한 이렇게 군집화한 유형 정보는 이탈 분류 모델에 활용하였다. 제안한 분류 모델은 6가지 알고리즘을 이용하여 구현했으며 다양한 측정 지표를 통해 성능을 평가하였다. 실험에 의하면 가장 높은 성능을 보인 알고리즘은 의사 결정 트리(C4.5)였으며 85.7%의 accuracy와 72.3%의 F-measure를 기록하였다.

본 연구에서 제안한 사용자 행동 유형 기반의 이탈 분석 기법은 이탈 예측에 활용될 수 있을 뿐만 아니라 이를 바탕으로 사용자 유형별 이탈 방지 정책이나 콘텐츠 개선 전략에 효과적으로 쓰일 수 있을 것으로 기대한다.

더 나아가 본 연구에서 제안한 행동 유형에 따른 이탈 분석 기법은 온라인 게임 분야뿐만 아니라 이탈 예측이 중요한 문제로 연구되고 있는 은행, 보험, 소매업, 통신업 등[1]에서도 응용 가능할 것으로 기대한다.

향후에는 본 연구에서 제안한 기법에 대해 다른 기간 및 게임 데이터를 이용하여 범용성을 검증할 예정이며, 피처를 더 보완함으로써 Recall을 높이는 개선 작업도 진행할 계획이다.

## References

- [1] J. Kawale, A. Pal, and J. Srivastava, "Churn Prediction in MMORPGs: A Social Influence Based an Approach," *Proc. of the International Conference Computational Science and Engineering, IEEE*, pp. 423-428, 2009.
- [2] K. Park, and M. Cha, "Churn Analysis of Maximum Level Users in Online Games," *Journal of KIISE*, Vol. 44, No. 3, pp. 314-322, Mar. 2017. (in Korean)
- [3] Z. H. Borbora, and J. Srivastava, "User Behavior Modelling Approach for Churn Prediction in Online Games," *Proc. of the 2012 International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2012 International Conference on Social Computing (SocialCom)*, IEEE, pp. 51-60, 2012.
- [4] F. Hadji, R. Sifa, A. Drachen, K. Thureau, K. Kers-ting, and C. Bauckhage, "Predicting Player Churn in the Wild," *Proc of the Conference on Computational Intelligence and Games (CIG)*, IEEE, 2014.
- [5] Z. Borbora, J. Srivastava, K.-W. Hsu, and D. Williams, "Churn Prediction in MMORPGs using Player Motivation Theories and an Ensemble Approach," *Proc. of the 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, IEEE, pp. 157-164, 2011.
- [6] J. Hadden, A. Tiwari, R. Roy, and D. Ruta, "Churn Prediction using Complaints Data," *Proc. of World*

*Academy Of Science, Engineering and Technology*, 2006.

- [7] K. B. Shores, Y. He, K. L. Swanenburg, R. Kraut, and J. Riedl, "The Identification of Deviance and its Impact on Retention in a Multiplayer Game," *Proc. of the 17th ACM conference on Computer supported cooperative work & social computing*, ACM, pp. 1356-1365, 2014.
- [8] T. Vafeiadis, K. I. Diamantaras, G. Sarigiannidis, and K. Ch. Chatzisavvas, "A comparison of machine learning techniques for customer churn prediction," *Simulation Modelling Practice and Theory*, Vol. 55, pp. 1-9, 2015.
- [9] K. Dasgupta, R. Singh, B. Viswanathan, D. Chakraborty, S. Mukherjee, A. A. Nanavati, and A. Joshi, "Social Ties and their Relevance to Churn in Mobile Telecom Networks," *Proc. of the 11th international conference on Extending database technology: Advances in database technology*, ACM, pp. 668-677, 2008.
- [10] N. Lu, H. Lin, J. Lu, and G. Zhang, "A Customer Churn Prediction Model in Telecom Industry Using Boosting," *IEEE Transactions on Industrial Informatics*, Vol. 10, No. 2, pp. 1659-1665, 2014.
- [11] 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," *Proc. of the 18th Asia Pacific Symposium on Intelligent and Evolutionary Systems*, Vol. 2, pp. 403-414, 2014.
- [12] J. Lee, J. Lim, W. Cho, and H. K. Kim, "I know what the BOTs did yesterday: Full action sequence analysis using Naïve Bayesian algorithm," *Proc. of the 2013 12th Annual Workshop on Network and Systems Support or Games (NetGames)*, IEEE, pp. 1-2, 2013.
- [13] J. Lee, S. W. Kang, and H. K. Kim, "A study on hard-core users and bots detection using classification of game character's growth type in online games," *Journal of Korea Institute of Information Security and Cryptology*, Vol. 25, No. 5, pp. 1077-1084, 2015.
- [14] H. J. Choi, J. Y. Woo, and H. K. Kim, "Online Game Identity Theft Detection Model based on Hacker's Behavior Analysis," *Journal of Korea Game Society*, Vol. 11, No. 6, pp. 81-98, 2011. (in Korean)
- [15] J. Yang, X. Wei, M. S. Ackerman, and L. A. Adamic, "Activity Lifespan: An Analysis of User Survival Patterns in Online Knowledge Sharing Communities," *Proc. of ICWSM*, pp. 186-193, 2010.



오 세 준

2017년 서울시립대학교 수학과 졸업(학사). 2017년~현재 고려대학교 정보보호대학원 재학(석사). 관심분야는 온라인 게임 보안, 데이터 마이닝



이 은 조

2002년 숭실대학교 컴퓨터학부 졸업(학사) 2007년 숭실대학교 정보통신대학원 졸업(석사). 2015년~현재 고려대학교 정보보호대학원 재학(박사). 2007년~현재 엔씨소프트 데이터분석모델링팀 팀장. 관심분야는 온라인 게임 보안, 데이터 마이닝



우 지 영

2000년 KAIST 산업공학과 졸업(학사) 2002년 KAIST 산업공학과 졸업(석사) 2006년 KAIST 산업공학과 졸업(박사) 2006년~2008년 삼성화재 CRM파트 과장. 2008년~2010년 미국 아리조나대학 인공지능연구소 연구원. 2011년~2016년 고려대학교 정보보호대학원 연구교수. 2016년~현재 순천향대학교 빅데이터공학과 교수. 관심분야는 데이터 마이닝



김 휘 강

1998년 KAIST 산업경영학과 졸업(학사). 2000년 KAIST 산업공학과 졸업(석사). 2009년 KAIST 산업및시스템공학과 졸업(박사). 2004년~2010년 엔씨소프트 정보보안실장, Technical Director. 2010년~2014년 고려대학교 정보보호대학원 조교수. 2015년~현재 고려대학교 정보보호대학원 부교수 관심분야는 온라인 게임 보안, 네트워크 보안, 네트워크 포렌식