

온라인 게임에서 사용자의 이탈 예측 및 이탈 사유 분석

소셜 활동 성향 및 게임 참여도를 중심으로

User Behavior Analysis for Predicting Churn of Loyal Customers in Online Games based on Social Relationships and Degree of Participation

| | |
|--------------------|---|
| 저자 (Authors) | 서은비, 우지영, 김휘강 Eunbi Seo, Jiyoung Woo, Huy Kang Kim |
| 출처 (Source) | 정보과학회논문지 45(11) , 2018.11, 1124-1134(11 pages) Journal of KIISE 45(11) , 2018.11, 1124-1134(11 pages) |
| 발행처 (Publisher) | 한국정보과학회 KOREA INFORMATION SCIENCE SOCIETY |
| URL | http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07556786 |
| APA Style | 서은비, 우지영, 김휘강 (2018). 온라인 게임에서 사용자의 이탈 예측 및 이탈 사유 분석. 정보과학회논문지, 45(11), 1124-1134 |
| 이용정보 (Accessed) | 한신대학교 218.155.***.152 2019/07/19 17:02 (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.

온라인 게임에서 사용자의 이탈 예측 및 이탈 사유 분석 : 소셜 활동 성향 및 게임 참여도를 중심으로

(User Behavior Analysis for Predicting Churn of
Loyal Customers in Online Games based on
Social Relationships and Degree of Participation)

서 은 비 [†]
(Eunbi Seo)

우 지 영 ^{††}
(Jiyoung Woo)

김 휘 강 ^{†††}
(Huy Kang Kim)

요 약 온라인 롤플레이 게임(MMORPG) 내 사용자들은 다양한 소셜 활동 성향을 보이며, 일부 사용자의 경우 혼자 게임을 즐기는 성향을 나타내기도 한다. 본 논문에서는 사용자가 소속된 길드의 특성을 소셜 활동 및 소속감 정도에 따라 분류하고, 분류된 각 그룹의 이탈율 및 이탈원인을 분석한다. 또한 소셜 활동 성향으로 분류된 각 사용자 그룹을 대상으로 게임 참여도 변동 추이를 측정하여 이탈을 예측하는 프레임워크를 제안한다. 비슷한 성향으로 분류된 각 그룹의 사용자는 이탈 직전에 유사한 행동 패턴을 보일 수 있으므로, 이를 기준으로 이탈 사용자와 비이탈 사용자의 패턴을 분류할 수 있다. 엔씨소프트의 대표 MMORPG인 아이온 데이터를 대상으로 본 모델을 테스트하였으며, 평균 약 75%의 정밀도를 보여주었다.

키워드: 유저 행위 분석, 이탈 사용자 예측, 온라인 롤플레이 게임, 데이터마이닝

Abstract Game users in MMORPGs engage in a variety of social activities. However, a few users tend to play games alone, and are designated 'loners' similar to modern society. We classified game guilds and game users based on similar user behaviors and community characteristics. We propose a model that predicts churn users by measuring the participation of users in each group. Users in each group show similar behavioral patterns, suggesting that we can classify churn users along with ordinary users. We tested this model for NCsoft's MMORPG, Aion. Using Randomforest, the recall was measured at an average of 75%.

Keywords: user behavior analysis, churn user prediction, MMORPG, data mining

[†] 학생회원 : 고려대학교 정보보호대학원 정보보호학과
dmsql0722@hksecurity.net

^{††} 정 회 원 : 순천향대학교 빅데이터공학과 교수
jywoo@sch.ac.kr

^{†††} 종신회원 : 고려대학교 정보보호대학원 정보보호학과 교수(Korea Univ.)
cenda@korea.ac.kr
(Corresponding author)

논문접수 : 2017년 10월 19일
(Received 19 October 2017)

논문수정 : 2018년 7월 12일
(Revised 12 July 2018)

심사완료 : 2018년 8월 9일
(Accepted 9 August 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.
정보과학회논문지 제45권 제11호(2018. 11)

1. 서론

MMORPG는 적게는 수십 명, 많게는 수천 명 이상의 사용자가 상호작용을 하며 다양한 게임 콘텐츠를 즐길 수 있는 RPG 게임이다. 연구 결과에 따르면 전 세계에서 인터넷 사용자의 약 40%가 여가 시간에 온라인 게임을 즐기며[1], 이 중 MMORPG의 사용자는 주당 평균 25.86시간을 게임에 소비하고 있다[2]. 이처럼 MMORPG의 인기가 증가함에 따라, 게임 유저의 확보 및 유지 전략에 대한 연구는 게임 산업에서 가장 중요한 관심사가 되었다. 또한 일반적으로 신규 고객의 확보는 기존 고객의 유지비용보다 더 높기 때문에[3-5], 기존 고객을 유지 및 관리하는 연구는 기업의 필요와 맞물려 여러 분야에서 활발히 연구 되고 있다[6-8].

인-게임 내 이탈 고객은 게임 콘텐츠 및 다른 게임 유저의 상호작용 과정에서 게임에 흥미를 잃게 되어 더 이상 접속을 하지 않는 사용자를 의미한다. 이탈 사용자의 증가는 게임 회사의 수익성에 막대한 영향을 끼칠 뿐만 아니라, 다른 사용자에게 부정적인 영향을 끼치며 이들의 이탈을 유도할 수 있기 때문에[9], 사용자의 이탈 원인 및 이를 예측하는 연구의 필요성이 대두되고 있다.

최근 컴퓨팅 기술의 발달로, 온라인 게임의 사용자 행위를 서버 단에 게임 로그의 형태로 기록할 수 있게 되었다. 더구나 MMORPG 내 가상 세계가 현실 세계와 유사하게 설계됨에 따라, 가상현실의 사용자들을 들여다봄으로써 현실 세계의 다양한 현상을 이해할 수 있게 되었다[10]. 이는 과거 게임 도메인의 이탈 고객 연구가 활발히 이루어지지 않았던 것과는 달리, 온라인 게임의 사용자 이탈 연구에 관심을 갖게 되는 계기가 되었다.

MMOPRG의 사용자는 사냥, 채집, 퀘스트 등 다양한 활동을 하며 다른 사용자와의 커뮤니티(파티, 거래)를 형성한다. 하지만 게임 내에서는 사용자 간 친목에 관심을 두지 않는, 혼자 게임을 즐기는 유형 또한 존재한다. 이와 같은 유형은 친목 도모를 선호하는 유형과 다른 게임 패턴을 가질 수 있으며, 따라서 사용자의 이탈에 영향을 끼치는 원인이 다를 가능성이 있다. 즉 가상 사회 내에서의 소셜 활동 및 사회적 관계는 사용자의 게임 몰입도 및 흥미에 영향을 미치며, 사용자는 소셜 활동의 성향에 따라 다른 이탈 원인 및 행동 패턴을 보일 수 있다.

기존의 연구에서는 사용자의 활동 시간 등 특정 행위를 분석하거나[9], 더 나아가 소셜 네트워크의 관점에서 사용자의 게임 참여도의 변동 추이를 분석하여 이탈을 예측하는 방안이 제시되었다[10]. 본 논문에서는 사용자의 소셜 활동 성향 및 게임 참여도에 중점을 두어 사용

자의 이탈 원인을 분석하고 이탈을 예측하는 모델을 제안한다. 제안하는 모델의 방법은 다음과 같다.

첫째, 길드를 길드 내 사용자의 소셜 활동 및 소속감에 따라 분류한다. 예컨대 커뮤니티를 원하지 않는 사용자라도 다른 목적(아이온에서는 길드 내 동시 접속자가 5명 이상일 경우 추가 경험치를 받을 수 있음)을 위해 길드에 소속될 가능성이 있다.

둘째, 길드 타입에 따라 분류된 각 그룹에 독립적으로 이탈 예측 알고리즘을 적용한다. 비슷한 소셜 활동 성향을 가진 사용자는 이탈의 원인도 유사할 수 있다. 각 그룹의 사용자는 이탈 직전까지 유사한 행동 패턴을 보일 수 있으므로, 이를 기준으로 사용자의 이탈 패턴과 비이탈 패턴을 분류한다. 이 때 유저의 게임 활동량의 변동률을 측정할 수 있는 피처를 입력 피처로 사용하며, 분석 시점을 기준으로 최근 게임 활동량 변동에 더 많은 가중치를 둔다.

본 논문에서 제안하는 이탈 예측 모델을 검증하기 위해, 엔씨소프트의 대표 MMORPG인 아이온의 데이터를 사용하였다. 본 논문의 구성은 다음과 같다. 제 2장에서는 기존 연구 결과를 검토하고, 이를 바탕으로 본 연구의 목표를 도출한다. 제 3장에서는 본 논문에서 제안한 모델의 주요 방법론을 설명한다. 제 4장에서는 데이터셋 정보 및 전처리 과정을 설명하고 제안한 모델을 적용하여 결과 및 정확도를 분석한다. 제 5장에서는 본 연구의 시사점에 대해 서술한다.

2. 관련 연구

서비스 기반의 산업에서 이탈 고객의 증가는 산업 매출에 직접적인 영향을 끼치므로, 고객의 행동 변화를 추적해 이탈 고객을 정확하게 예측하는 방법은 전반적인 산업에서 활발하게 연구되고 있다. 전통적으로 고객 이탈 분석은 통신 산업에서 가장 활발히 분석되어졌다. 또한 소매업, 은행, 보험, 네트워크 및 최근 온라인 게임 장르 중 하나인 MMORPG 내에서도 고객 이탈 예측에 대한 다양한 기술이 발전하고 있다.

고객 이탈 분석 연구에서는 로지스틱 회귀 분석, 의사결정 트리, 신경망 네트워크 등 다양한 머신 러닝 모델이 사용된다. Datta등은 휴대 전화 고객을 대상으로 고객의 불만 사항 데이터에 의사 결정 트리 알고리즘을 적용해 이탈 고객을 예측하였다[11].

Yaya등은 IBRF(Equalized Balanced Random Forest)이라 불리는 새로운 학습 방법을 구현하여 은행의 고객을 유지하는 모델을 제안하였다[12]. IBRF는 각 클래스의 분포를 변경하며 적절한 피처를 반복적으로 학습시키고, 소수 클래스의 오차에 높은 패널티를 부과한다. 해당 모델은 실제 은행 고객의 데이터에 적용되었으며,

인공 신경, 의사결정트리 및 SVM과 같은 알고리즘과 비교하였을 때 예측 정확도를 크게 향상시키는 것으로 증명되었다.

Chih-Fong등은 역 전파 인공 신경망(ANN)과 자가 조직 지도 신경망(SOM)의 두 가지 신경망 네트워크 기술을 결합한 하이브리드 모델을 제안하였다[13]. 하이브리드 모델은 ANN 및 SOM이 결합된 모델로, 여러 도메인을 대상으로 단일 신경망 모델 및 ANN의 기본 모델보다 우수한 성능을 보여주었다.

Bingquan등은 4개월간의 통화 관련 데이터(회선 정보, 청구서 및 지불 정보, 계좌 정보, 인구 통계, 불만 정보 등)를 이용하여 고객 이탈 예측을 위한 연구를 수행하였다[14]. 새로운 피처에 대해 Navies Bayes, SVM, 회귀 분석 등 7가지의 새로운 예측 기법을 적용하였으며, 비교 실험을 통해 고객 이탈 예측을 위한 7가지 모델링 기법을 평가하였다. 실험 결과, 이 중 6가지 모델링 기법은 통신 서비스 분야 내 고객 이탈 예측을 위한 기존의 연구와 비교하여 더 효과적인 연구 결과를 도출하였다.

Zhao등은 개선된 SVM 알고리즘을 이용한 고객 이탈 예측 모델을 제안하였으며, 무선 산업 고객 데이터를 대상으로 모델을 테스트하였다[15].

이처럼 고객 예측 모델에 대한 연구는 다양한 도메인을 기반으로 발전하고 있으며, 온라인 게임 산업 내에서도 이익 창출을 위해 고객 유지에 대한 활발한 연구가 이루어지고 있다.

유저 간 커뮤니티 및 소셜 활동이 활발한 온라인 게임 장르인 MMORPG는 소셜 네트워크를 이용한 이탈 예측 모델의 연구가 활발히 진행되었다. 한 사례로 Jaya등은 게임 유저의 사회적 영향(긍정적, 부정적 요소)에 따른 분류 및 게임 참여도의 측정을 통해 이탈을 예측하는 모델을 제안하였다[9]. Jaya등은 사용자의 사회적 영향을 벡터로 표현하여 사용자의 게임 참여도와 결합하였으며, 기존의 모델과 비교하여 예측 정확도를 향상시켰다.

Zoheb등은 사회과학적 관점에서의 이탈 예측을 연구하고, 사용자의 동기 이론을 통합한 예측 모델을 제안하였다[16]. 이론 중심의 접근법은 데이터 기반 모델보다 정확도는 떨어지지만 도메인 전문가가 해석하기에 더 용이하며, 고객 수가 제한되어 있는 경우 더 많은 고객 식별을 통해 투자 수익률을 높일 수 있음을 증명하였다. 또한 클러스터링 기술을 사용하여 데이터를 분류한 후, 분류된 각 데이터의 양상불을 생성하여 성능을 향상시켰다. 실험 결과 해당 모델은 리콜 값에서 기존 분류 방법보다 훨씬 우수한 성능을 보였다.

Julian등은 명시적으로 사용자의 시간적 흐름에 따른

변동을 다룰 수 있는 Markov 모델을 구현하였으며, 일반적인 소셜 게임을 대상으로 예측 성능을 비교하였다[17]. 제안한 모델의 신경망에 AUC 영역에서 최상의 예측 성능을 보여주는 것을 증명하였다. 또한 이탈 직전의 사용자에게 인센티브를 주었을 때, 해당 사용자와의 의사소통에 대한 효율성이 크게 향상되는 것을 보여주었다.

기존 연구 결과와 같이 게임 내 사용자의 사회적 관계와 소셜 활동 성향은 사용자의 이탈 원인에 영향을 준다. 따라서 본 논문에서는 게임 내 소셜 활동 및 게임 참여도에 따라 비슷한 성향의 사용자 집합을 구축하여 분류 알고리즘을 적용함으로써, 신뢰도 높은 이탈 예측 모델을 구현한다. 모델의 성능을 높이기 위해 피처를 다양한 관점에서 선정하였으며, 사용자의 게임 활동량 측정 시, 분석 시점을 기준으로 최근 변화 값에 더 큰 가중치를 주었다. 본 논문은 사용자의 소셜 활동 성향을 기준으로 이탈 원인 분석 및 이탈 예측 모델을 구현하는 것을 목표로 한다.

3. 방법론

3.1 데이터 셋 및 전처리

엔씨소프트사의 아이온 게임 로그 데이터 셋을 실험 데이터로 사용하였다. Fig. 1과 Fig. 2는 분석 대상의 주 당 로그인 횟수 및 플레이 시간에 대한 누적분포함수이다. 분석 기간 내 해당 주에 한 번도 접속하지 않은 사용자는 약 54%로, 전체 사용자 중 매우 큰 비율을 차지하였다. 이러한 사용자는 게임 계정은 존재하지만 게임 활동을 하지 않는 사용자로, 이탈 가능성이 있는 휴면계정이라고 볼 수 있다. 즉 MMORPG의 이탈 기준은 '계정 탈퇴'가 아닌 일정 기간 이상 게임에 접속을 하지 않는 유저라 정의할 수 있다. 탈퇴 하지 않은 모든 휴면계정은 복귀 가능성이 존재하므로 본 논문에서는 이탈 기준을 분석 시점으로부터 13주 이상 한 번도 접속하지 않은 사용자로 정의하였다.

Fig. 3은 분석 기간 동안 사용자의 라이프 사이클을 분류한 표이다. 전체 데이터 셋의 기간은 2015년 12월 9일부터 2016년 6월 24일까지의 29주이며, 해당 기간 동안의 전체 사용자 수는 149,743명이다. 이탈 분석 시점은 17주차로 선정하였으며, 따라서 사용자의 행동 변화를 분석하는 기간은 17주차 전까지인 16주이다. 본 논문에서는 각 주 별 사용자의 행동 변화를 관찰하므로, 해당 주에 접속한 사용자만을 추출하였다. 또한 데이터 셋을 제공한 엔씨소프트사에서 사용하는 것과 같이 사용자를 일반 사용자와 봇 사용자로 분류하였다. 게임 봇 사용자의 경우 이탈 방지 대상이 아니므로, 게임 수익에 직접적인 영향을 주는 일반 사용자를 대상으로 이탈 원

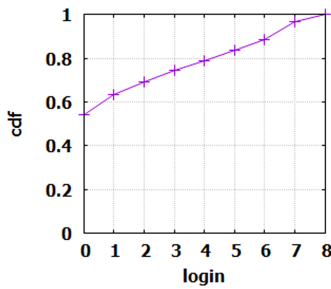


그림 1 로그인 누적 분포 함수

Fig. 1 Login cumulative distribution function

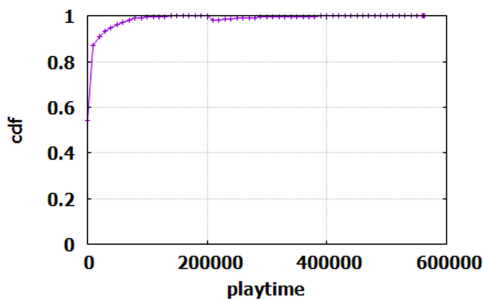


그림 2 평균 플레이 시간 누적 분포 함수

Fig. 2 Playtime cumulative distribution function

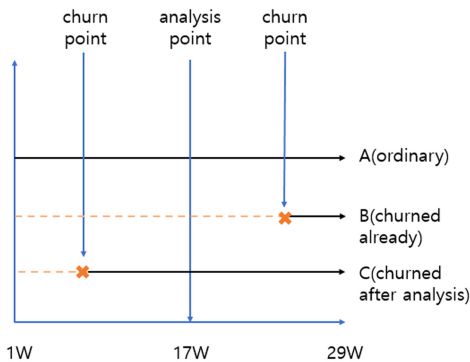


그림 3 사용자 라이프 사이클 분류

Fig. 3 User life cycle

인을 파악하고 이를 예측하였다. 즉 분석 기간 동안 길드에 속한 적이 있는 사용자의 수는 16,460명이며, 이 중 각 주를 기준으로 한번 이상 접속한 유저는 10,470명이다. A 그룹은 분석 기간 동안 꾸준히 접속한 사용자, B 그룹은 분석 시점 이전에 이미 이탈을 한 사용자, 마지막으로 C 그룹은 분석 시점 이후 이탈한 사용자이다. 본 논문에서는 분석 시점 이전에 이탈한 사용자 그룹을 제외한 A와 C에 속한 사용자만을 연구 대상으로 삼았다. 최종적으로 분석 시점 이전에 이탈한 B 그룹을

제외한 분석 대상의 전체 사용자 수는 9,637명이다. 이탈 사용자로 분류된 C 그룹의 경우 3,896명이며 향후 이탈 예측 모델의 정답지로 사용하였다. 전처리 결과 분석 대상의 전체 사용자 중 약 40%가 이탈함을 보였다.

3.2 길드 분류

MMORPG 게임의 '길드'는 게임 중 특정 목적을 위해 모인 대규모 커뮤니티로서, 파티를 비롯한 다양한 소셜 활동을 지원한다. 따라서 게임 유저들은 특정 길드에 가입하여 다른 유저와의 커뮤니티를 도모하는 경우가 많다. 접속자의 수가 5명 이상일 경우, 추가적인 경험치를 제공하는 이벤트를 진행하고 있다. 다른 유저와의 소셜 활동을 위해 길드에 가입하는 경우가 일반적이지만, 어딘가에 소속되거나 소셜 활동을 하는 것을 좋아하지 않는 이른바 '혼자 게임을 즐기는 유저'는 경험치 버프용, 즉 경험치만을 얻기 위한 목적 등 친목 외 다른 목적을 위해 길드에 가입할 가능성이 높다. 이러한 유저의 경우 길드의 이동이 잦거나, 길드에 대한 소속감 및 길드 활동을 보이지 않는 경향을 보일 수 있다.

따라서 본 연구에서는 아래와 같은 피처를 이용해 K-means 클러스터링 알고리즘을 수행하였다. 피처는 길드 내 사용자의 활동을 볼 수 있는 피처로 선정하였으며, 이를 통해 길드원간의 소속감 및 소셜 활동에 따라 길드를 분류할 수 있다.

- concurrency: 최대 동시 접속자 수
- guild join: 해당 길드 가입 횟수
- guild out: 해당 길드 탈퇴 횟수
- partyplay with guild: 길드원간 파티 횟수
- partyplay with other: 길드 밖 파티 횟수
- friend in guild: 길드 내 친구 수
- friend not in guild: 길드 밖 친구 수
- average level: 평균 레벨
- average playtime: 평균 플레이 시간
- average experience: 평균 경험치
- join with invite: 초대로 가입한 횟수
- join with no approval: 승인 없이 가입한 횟수
- guild money: 길드 자금
- total user: 길드원 수

최근 연구 결과에서는 가상 커뮤니티에서의 구성원 행동을 설명함에 있어 소속감을 가장 중요한 요소로 강조한다[18]. 구성원의 소속감은 가상 커뮤니티의 성공을 설명할 주요 변수이자, 가상 커뮤니티 구성원들의 행동을 설명하는 주요 요인으로 부각되고 있다. 따라서 본 논문에서는 아이온 내 가상 커뮤니티의 성향이 사용자의 소속감에 높은 영향을 미칠 수 있다고 가정하고 사용자를 소속된 길드의 성향, 즉 앞서 분류한 길드 타입에 따라 분류하였다. 이 때 사용자가 속해 있는 길드를

주 단위로 추출하여 분석 시점과 가장 가까운 주에 속해 있는 길드를 해당 사용자의 대표 소속 길드로 정의하였다.

3.3 이탈 예측 모델

사용자의 소속 길드의 성향 및 소셜 활동 성향에 따라 이탈 직전에 보이는 행동 패턴이 다를 수 있다. 따라서 소속된 길드의 성향에 따라 분류된 각 그룹 내 사용자의 이탈 비율을 비교하여 소셜 활동 성향과 이탈 원인과의 상관관계를 분석하고, 이탈 예측 모델을 구현하였다.

Fig. 4는 본 논문에서 제안한 이탈 유저 분석 모델 과정이다. 길드 분류를 통해 사용자를 소셜 활동 성향에 따라 그룹화하고, 각 그룹에 독립적으로 사용자의 다양한 게임 활동 변동률을 측정하여 이탈을 예측하였다. 이때 측정되는 게임 활동의 종류의 경우, 사용자의 게임 참여도를 볼 수 있는 아래와 같은 피처로 선정하였다. 즉 아이온 내 사용자가 보여주는 대표 게임 행위를 기준으로 피처를 선정하였다.

- teleport: 텔레포트 횟수
- que_suc: 퀘스트 성공 횟수
- playtime: 평균 플레이 시간
- login: 로그인 횟수
- indun : 인턴 입장 횟수
- dice: 주사위 횟수
- harvest_item: 아이템 수집 횟수
- glide: 글라이드(비행) 횟수
- combine_item: 아이템 강화 횟수
- pvp : pvp 횟수
- partyplay : 파티플레이 횟수

3.2에서 언급한 길드의 소속감을 나타내는 피처들을 이용해 K-means 클러스터링을 수행하여 길드를 분류하고, 사용자를 소속된 각 길드의 소셜 활동 성향에 따

라 분류하였다. 분류된 각 사용자 그룹에 대해 다양한 게임 활동에 대한 추세, 변동률을 측정한 후, 측정된 값을 기계 학습 모델의 학습 피처로 입력하여 최종적으로 이탈 예측 알고리즘을 수행하였다.

Borbora등은 사용자의 게임 참여도를 측정하기 위해 enthusiasm과 persistence라는 2가지 개념의 변수를 제안하였다[10]. 실험 결과 이를 적용한 이탈 예측 모델은 동일한 데이터를 대상으로 수행하였던 이전 모델보다 더 높은 정확도를 보였으며, 본 논문에서는 Borbora가 제안한 enthusiasm과 persistence의 변수를 이탈 예측 모델의 입력 피처로 활용하였다.

3.3.1 Enthusiasm

연속 주 동안 관찰된 행동의 증가 또는 감소의 양을 기울기의 함으로 측정하였다. Fig. 5에서 X_i 는 해당 주에 관찰된 행동의 빈도수를 나타내며, X_{i-1} 은 해당 주의 바로 전 주에 관찰된 행동의 빈도수를 나타낸다. N 은 분석기간 동안 주의 총 수이며, W_i 는 분석 기간 중 해당 주가 몇 번째 주인지를 의미한다. W_i 를 곱하는 선형 가중 함수를 사용하여 최근 변화가 더 많은 비중을 차지하도록 하였다. 또한 해당 주와 직전 주의 빈도수 차를 곱하여 행동의 변동률을 측정하였다.

$$x_{enthu} = \sum_{i=1}^N w_i * (x_i - x_{i-1})$$

그림 5 Enthusiasm 알고리즘

Fig. 5 Enthusiasm algorithm

3.3.2 Persistence

Enthusiasm의 개념과 유사하지만, 본 개념의 경우 연속 주 동안 얼마나 꾸준히 특정 활동량이 변하였는지

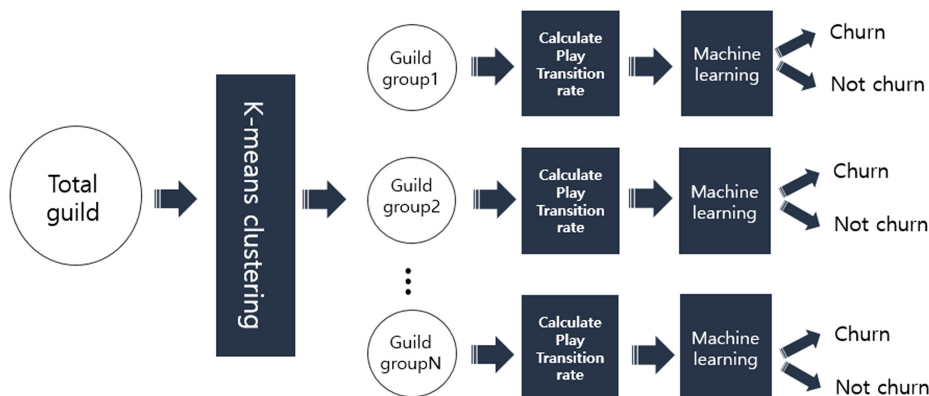


그림 4 이탈 예측 모델 과정

Fig. 4 Churn prediction process

$$x_{persist} = \sum_{i=1}^N w_i * Ind(x_i - x_{i-1})$$

$$Ind(x_i - x_{i-1}) = \begin{cases} 1, & \text{for } x_i - x_{i-1} > 0 \\ 0, & \text{for } x_i - x_{i-1} = 0 \\ -1, & \text{for } x_i - x_{i-1} < 0 \end{cases}$$

그림 6 Persistence 알고리즘
Fig. 6 Persistence algorithm

해당 변화의 흐름을 계산하는 것에 더 중점을 두는 개념이다. 사용자의 게임 활동에 대한 일반적인 분위기를 방향(위/아래)의 변화로 표현하여 측정하였다. Fig. 6에서 보듯이, Persistence는 Enthusiasm과 달리 각 행동의 변화율을 1, 0, -1로 나누는 indicator function을 사용하였으며, 최근 변화에 더 많은 가중치를 두었다.

이렇게 구해진 사용자의 다양한 게임 활동들의 변동률을 기계 학습 알고리즘의 이탈 예측 피처로 입력하였다. 따라서 각 게임 활동은 Enthusiasm 값과 Persistence 값의 두 종류의 변동률 값을 가지게 된다. 기계 학습 알고리즘을 통해 트레이닝 데이터 셋 내 비슷한 성향을 가진 사용자 간 게임 활동 추세를 비교하여, 최종적으로 테스트 셋 내 사용자를 이탈과 비이탈로 분류할 수 있다.

4. 실험 결과 및 분석

4.1 길드 분석

먼저 사용자가 소속된 각 길드의 성향을 정의하기 위해 K-means 클러스터링을 수행하였다. Fig. 7은 전체 제공 합을, Fig. 8은 군집 간 제공 합을 나타낸다. Fig. 7에서 클러스터 수, k 값이 작아짐에 따라 전체 제공 합이 감소하는 것을 볼 수 있다. k가 3 이상일 경우 감소 비율이 급격하게 떨어지므로, k-means 클러스터링의 적정 k 값을 3으로 선정하였다.

길드 분류에 사용된 피처는 3.2에서 제시한 길드원 간 커뮤니티 및 길드 내 활동을 볼 수 있는 피처로 구성하

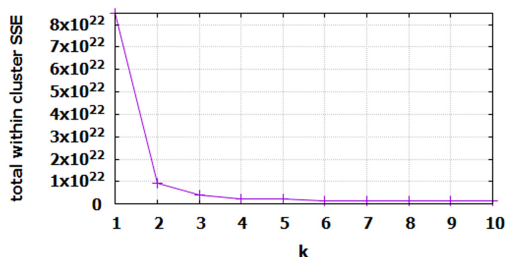


그림 7 전체 제공합
Fig. 7 Total within-cluster SSE

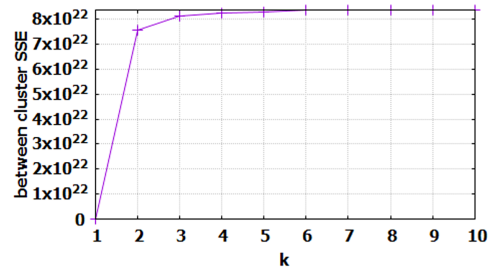


그림 8 군집 간 제공 합
Fig. 8 Between-cluster SSE

며, 이 중 중요한 영향을 끼치는 변수를 알기 위해 feature selection 알고리즘 중 하나인 CfsSubsetEval를 수행한 결과 길드원의 평균 레벨, 길드 가입 경로가 중요 피처로 선정되었다. 또한 각 피처는 0~1 사이의 값으로 정규화를 수행하였다. 길드 분석 결과 Table 1과 같이 길드 내 활동이 있는 활성화 길드와 길드 내 활동이 거의 없는 비활성화 길드로 분류되었으며, 활성화 길드 내에서는 친목 위주의 길드와 개인플레이가 위주인 길드로 분류되었다. 최종적으로 분류된 길드 그룹은 활성화 길드 내 개인플레이 길드(solo-play guild)와 친목 도모용 길드(co-play guild), 그리고 비활성화 길드(inactive guild)이다.

4.1.1 개인플레이 길드(solo-play guild)

분석 결과 분류된 길드 그룹 1은 길드 활동이 있는 활성화 길드로 Fig. 9에서 볼 수 있듯이 길드원이 길드를 가입하거나 탈퇴하는 경우가 높았으며, 이는 길드 내 소속감이 낮은 것을 의미한다. 평균 플레이 시간 및 길드 자금이 가장 높은 길드 그룹 2와 달리, 길드 그룹 1의 경우 전체 길드원의 수가 가장 높은 반면 길드 자금이 가장 적었다. 또한 Fig. 10과 같이 길드 내 최대 동시 접속자 수가 가장 많은 경향을 보았을 때 길드 그룹 1은 동시 접속자 수에 의한 경험치 버프용 목적 등으로 생성된 길드라 볼 수 있다. 따라서 길드 그룹 1을 친목 위주가 아닌 자기 다른 목적을 갖고 가입한 경우가 많은 개인플레이 길드로 분류하였다.

표 1 길드의 K-means 클러스터링 결과(k=3)
Table 1 K-means clustering result of guild(k=3)

| guild type | | explanation | # |
|----------------|-----------------|---|-------|
| active guild | solo-play guild | The guild that shows individual play guild for other purposes | 607 |
| | co-play guild | The guild that shows active social activities | 3,831 |
| inactive guild | inactive guild | The guild that show little activity | 1,916 |

4.1.2 친목 도모용 길드(co-play guild)

길드 그룹 2 또한 길드 활동이 있는 활성화 길드로, Fig. 9와 Fig. 11에서 볼 수 있듯이 길드 내 친구 수, 길드원과의 파티 플레이 횟수가 가장 높았으며, 길드원과의 커뮤니티가 활발한 것으로 보았을 때 친목 도모용 길드로 볼 수 있다. 반면, 앞서 정의한 개인플레이 길드는 길드원이 아닌 다른 친구들과 더 많은 교류를 하는 경향을 보였다. 따라서 길드 그룹 2를 친목 위주의 목적을 가진 친목 도모용 길드로 분류하였다.

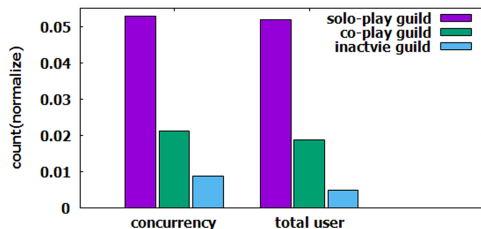


그림 9 길드 간 최대 동시 접속자 수 비교(정규화)
Fig. 9 The number of concurrency(normalization)

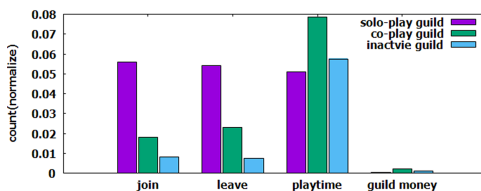


그림 10 길드 특성 비교(정규화)
Fig. 10 Guild characteristics(normalization)

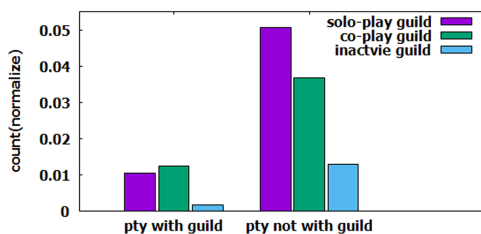


그림 11 길드 간 파티 플레이 횟수 비교(정규화)
Fig. 11 The number of concurrency(normalization)

4.1.3 비활성화 길드(inactive guild)

길드 그룹 3의 경우, 파티 플레이 활동, 친구 수, 평균 플레이 시간, 길드원 수 등 전반적인 부분에서 낮은 커뮤니티를 보이므로 활동이 거의 없는 비활성화 길드로 분류하였다.

4.1.4 길드 가입 경로

길드 분류에 가장 크게 영향을 미치는 피처는 유저가 길드에 가입하게 된 경로로, 사용자의 길드 가입 경로가

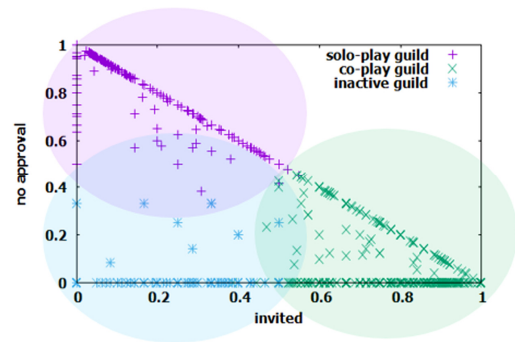


그림 12 길드 가입 유형
Fig. 12 Guild joining type

길드 내 소속감에 큰 영향을 끼치는 것을 알 수 있었다. 소속감이 높은 친목 도모용 길드의 경우 각 길드원끼리의 초대로 가입한 경우가 높았다. 반면 소속감이 낮은 개인플레이 길드의 경우 길드의 가입/탈퇴가 잦으며, 길드에 승인 없이 가입할 수 있었던 경우가 높았다. Fig. 12는 분류된 각 길드 그룹의 길드 가입 경로 분포도로, 초대로 가입한 경우와 승인 없이 가입한 경우의 횟수는 서로 반비례하는 경향을 볼 수 있다

4.2 이탈자 분석

Table 2는 분류된 각 그룹에 속해 있는 사용자의 이탈 비율을 보여준다. 개인플레이 길드의 경우 가장 큰 이탈 비율을 보인 반면, 친목 도모용 길드의 경우 가장 낮은 이탈 비율을 보였다. 또한 전반적으로 게임 활동 및 길드 내 활동이 없었던 비활성화 길드 그룹의 경우 약 43%의 비율로 개인플레이 그룹보다 더 낮은 이탈 비율을 보였다. 즉 소셜 활동 및 소속감이 높은 사용자 그룹이 그렇지 않은 그룹보다 이탈할 확률이 더 낮은 것을 알 수 있다.

Table 3은 각 그룹 내 이탈자의 게임 활동에 대한 변동률(Persistence)의 평균을 비교한 표이며, 이를 Fig. 13의 그래프로 나타내었다. 양수 값의 경우에는 분석 기간 동안 활동이 증가함을, 음수 값의 경우에는 활동이 적어짐을 의미한다. 즉 0에 가까울수록 해당 활동에 대한 추세가 변하지 않음을 의미한다.

표 2 그룹에 따른 이탈 비율

Table 2 Churn rates according to user groups

| guild type | user group | churn user | total user |
|-----------------|-----------------|------------|------------|
| solo-play guild | solo-play group | 1,083(50%) | 2,172 |
| co-play guild | co-play group | 1,719(35%) | 4,920 |
| inactive guild | inactive group | 1,094(43%) | 2,545 |

표 3 그룹에 따른 이탈자의 persistence 수치

Table 3 Persistence value of churners in each user group

| activity | solo-play guild | co-play guild | inactive guild |
|-----------|-----------------|---------------|----------------|
| teleport | -0.28 | -1.51 | -1.95 |
| indun | 0.26 | -1.05 | -1.27 |
| glide | 0.26 | -1.01 | -1.53 |
| login | 1.99 | 1.20 | 0.01 |
| playtime | 0.66 | -0.67 | -1.56 |
| combine | -0.13 | -0.26 | -0.24 |
| harvest | 0.04 | -0.09 | -0.11 |
| dice | -0.02 | -0.07 | -0.29 |
| pvp | 0.06 | -0.04 | -0.42 |
| partyplay | -0.19 | -0.70 | -0.98 |

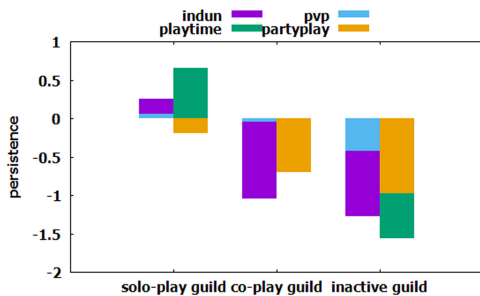


그림 13 그룹에 따른 이탈자의 persistence 수치

Fig. 13 Persistence value of churners in each user group

Fig. 13과 같이, 모든 사용자 그룹에서 이탈 직전의 기간 동안 전반적으로 게임 활동량이 감소하는 경향을 보였다. 그러나 파티 플레이 횟수의 경우 개인플레이 선호 그룹은 이탈 직전 큰 변화가 없었던 반면, 친목 도모 선호 그룹과 비활동 그룹은 비교적 큰 감소량을 보였다. 또한 개인플레이 선호 그룹은 pvp 횟수, 아이템 채집 횟수, 플레이시간, 클라이더 횟수, 인던 참여 횟수 등 몇

몇 게임 활동에 대해서 증가하거나 변화가 없는 경향을 보였다. 즉 이탈 직전 게임 활동량이 감소하는 경향을 보였던 다른 그룹들과 달리, 개인플레이 선호 그룹은 게임 활동량의 변화가 크지 않거나 되려 증가하는, 확연히 다른 경향을 보였다.

이탈 직전 게임 활동에 적극적이며 소셜 활동 및 소속감이 낮은 개인플레이 선호 그룹의 이탈 비율이 가장 높은 것으로 보아, 가상 사회 내 소셜 활동 및 소속감이 사용자의 이탈 여부에 영향을 미칠 수 있다는 것을 추측할 수 있다.

Table 5는 전체 사용자 중 이탈자와 비이탈자의 게임 활동에 대한 변동률(Persistence)의 평균을 비교한 표이며, 이를 Fig. 14의 그래프로 나타내었다. 이탈자는 양의 변동률을 보이는 비이탈자와 비교하여 음의 변동률을 보이며, 이탈 직전 인던, 파티플레이, PVP 등 게임 활동량이 더 감소하였다. 즉 이탈 예측 모델에 사용된 사용자의 행동 피쳐들은 전체 사용자의 이탈자 및 소셜 활동에 따른 각 그룹 내 이탈자의 행동 패턴을 잘 나타내는 것을 볼 수 있다.

표 5 전체 사용자의 persistence 수치

Table 5 Persistence value of total user

| activity | not churn user | churn user |
|-----------|----------------|------------|
| teleport | -0.06 | -1.29 |
| indun | 1.30 | -0.75 |
| glide | -0.08 | -0.80 |
| login | 4.71 | 1.09 |
| playtime | 0.61 | -0.56 |
| combine | -0.13 | -0.22 |
| harvest | -0.07 | -0.06 |
| dice | 1.20 | -0.12 |
| pvp | 0.35 | -0.12 |
| partyplay | 1.50 | -0.63 |

표 4 각 그룹의 이탈 예측 결과

Table 4 Prediction result for each user group

| group | algorithm | TPR | FPR | Precision | F-Measure | ROC Area | Recall |
|-----------------|---------------|------|------|-----------|-----------|----------|--------|
| solo-play guild | Random Forest | 0.77 | 0.23 | 0.78 | 0.77 | 0.86 | 0.77 |
| | PART | 0.82 | 0.43 | 0.66 | 0.73 | 0.72 | 0.82 |
| | Bagging | 0.79 | 0.25 | 0.76 | 0.78 | 0.84 | 0.79 |
| co-play guild | Random Forest | 0.64 | 0.14 | 0.72 | 0.64 | 0.84 | 0.64 |
| | PART | 0.79 | 0.29 | 0.60 | 0.68 | 0.77 | 0.79 |
| | Bagging | 0.63 | 0.18 | 0.66 | 0.64 | 0.82 | 0.63 |
| inactive guild | Random Forest | 0.74 | 0.19 | 0.74 | 0.74 | 0.85 | 0.74 |
| | PART | 0.79 | 0.39 | 0.61 | 0.69 | 0.75 | 0.79 |
| | Bagging | 0.71 | 0.23 | 0.70 | 0.71 | 0.82 | 0.71 |
| total group | Random Forest | 0.62 | 0.20 | 0.67 | 0.64 | 0.8 | 0.62 |
| | PART | 0.67 | 0.26 | 0.63 | 0.64 | 0.77 | 0.66 |
| | Bagging | 0.65 | 0.21 | 0.67 | 0.66 | 0.79 | 0.68 |

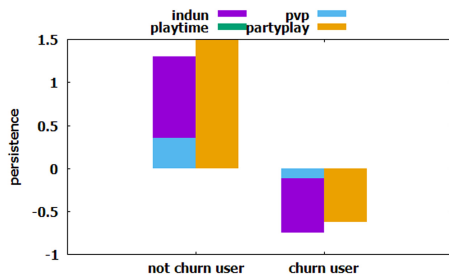


표 14 전체 사용자의 persistence 수치
Fig. 14 Persistence value of total user

4.3 이탈 예측

4.3.1 이탈 예측 평가 기준

이탈 예측에 대한 평가 기준은 Table 4와 같이 구성된다. TP(True Positive)는 이탈 사용자를 이탈 사용자로 분류된 횟수를, FP(False Positive)는 일반 사용자를 이탈 사용자로 잘못 분류한 횟수를 의미한다. 또한 FN(False Negative)은 이탈 사용자를 일반 사용자로 분류한 횟수를, TN(True negative)은 일반 사용자를 일반 사용자로 분류한 횟수를 의미한다. ROC area는 ROC curve의 밀면적으로 1에 가까울수록 좋은 성능을 보여준다. 아래는 각 평가 기준에 대한 수식을 보여준다.

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

수식 1 precision function

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

수식 2 recall function

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

수식 3 Accuracy function

$$F-measure = \frac{Precision * recall}{Precision + recall} * 2$$

수식 4 F-measure function

4.3.2 이탈 예측 피쳐 값

앞서 이탈 예측을 위해 Borbara의 2가지 enthusiasm과 persistence라는 변동을 측정 개념을 사용하였다. 이탈 예측 결과 enthusiasm을 통해 계산된 피쳐 값들의 경우 Persistence와 비교하여 더 낮은 정확도를 보여주었다. 따라서 가장 높은 이탈 예측율을 보여준 Persistence로 측정된 변동을 피쳐 값만을 본 모델의 이탈 예측 피쳐로 사용한다. 즉 3.3에서 설명한 이탈 예측 피쳐들의 각 persistence값을 구한 피쳐 목록들로 구성된다.

표 6 이탈 예측 피쳐 종류

Table 6 Prediction feature

| prediction feature |
|--------------------|
| teleport_per |
| que_suc_per |
| playtime_per |
| login_per |
| indun_per |
| dice_per |
| harvest_item_per |
| glide_per |
| combine_item_per |
| pvp_per |
| partyplay_per |

최종적으로 선정된 이탈 예측 피쳐는 Table 6과 같다.

4.3.3 이탈 예측 알고리즘

본 연구에서는 사용자를 소속된 길드의 소셜 활동 성향에 따라 분류하여, 각 그룹에 독립적으로 이탈 예측 모델을 적용하였다. 비슷한 성향으로 분류된 각 그룹의 사용자는 이탈 직전에 유사한 행동 패턴을 보였으며, 이를 기준으로 이탈 사용자와 비 이탈 사용자의 패턴을 분류하였다. 즉 로그 단의 게임 데이터에서 이탈 예측을 위한 유용한 피쳐들로 가공하고, 사용자의 소셜 성향에 따라 독립적으로 예측 알고리즘을 적용할 수 있는 전체적인 프레임 워크를 제안하였다.

본 모델에서 사용된 이탈 예측 알고리즘은 기존 연구에서 효율성이 검증된 알고리즘으로, Random Forest, PART, Bagging 알고리즘으로 구성된다. meta는 상황에 따라 적절한 알고리즘을 선택 및 조합하여 학습하는 선택적 메타학습으로, meta의 대표 알고리즘인 bagging (Bootstrap aggregating)은 분산을 줄이기 위한 알고리즘이며 기본 학습자에 따라 분류 및 회귀를 수행할 수 있다[19]. tree 기반의 알고리즘인 Random Forest는 분류, 회귀 분석 등에 사용되는 앙상블 학습 방법의 일종으로, 앞서 설명한 Bagging 계열 중 가장 예측력이 좋은 알고리즘이라 볼 수 있다[20]. PART는 rule 기반의 알고리즘으로, C4.5 결정 트리를 반복적으로 적용하여 “최상의 리프 노드”를 구축한다[21].

tree 및 rule은 블랙박스 모델인 인공신경망과 달리 모델의 추리가 가능한 화이트 박스 모델을 사용한다. 이는 분석가 관점에서 실험 결과를 해석하기에 용이하며, 따라서 본 논문에서는 이탈자의 예측에 대한 원인 및 행위 변화를 관찰할 수 있도록 tree 및 rule 기반의 알고리즘을 분류 알고리즘으로 선정하였다.

4.3.4 이탈 예측 평가

Table 4는 전체 그룹과 분류된 각 사용자 그룹에 대한 정확도를 보여주는 표이며, Fig. 15는 정확도 측정

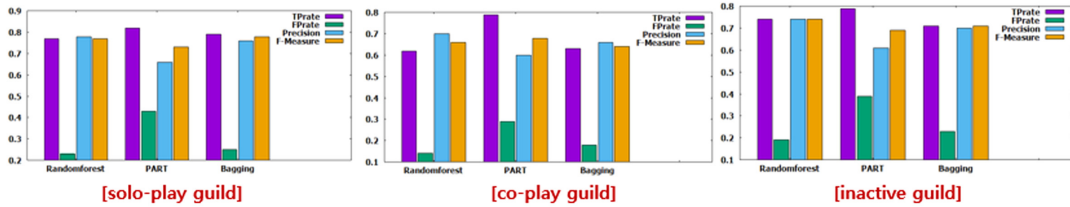


그림 15 이탈 예측 모델 성능 측정

Fig. 15 Churn prediction model performance

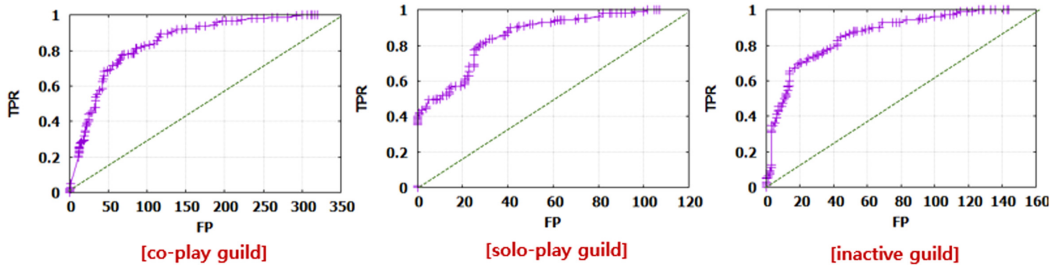


그림 16 Random Forest 사용 시 ROC curve 결과

Fig. 16 ROC curve using Random Forest

표 7 2017 게임 데이터마이닝 챌린지 F-정확도
Table 7 F-measure of 2017 datamining challenge

| Rank | Team | F-measure score |
|------|---------------|-----------------|
| 1 | Yokozuna Data | 0.62 |
| 2 | UTU | 0.60 |
| 3 | TripleS | 0.60 |
| 4 | TheCowKing | 0.60 |
| 5 | goedleio | 0.58 |

결과를 알고리즘 별로 비교한 그래프이다. Table 4의 total guild는 그룹을 분류하지 않고 이탈 예측 모델을 적용한 결과로, 그룹을 분류한 후 이탈 예측 모델을 적용한 결과보다 확연히 낮은 정확도를 보였다. 또한 그룹을 분류 후 이탈 예측 모델을 적용하였을 때, PART 알고리즘 이용 시 평균 0.8의 높은 TPR를 보였으며, Random Forest 알고리즘 이용 시 평균 0.19의 낮은 FPR를 보였다.

Fig. 16은 분류된 그룹을 대상으로 Random Forest 알고리즘을 적용한 이탈 예측 모델의 ROC curve이다. ROC curve는 X축인 1-특이도, Y축인 민감도를 통해 구축한 모델의 성능을 판단한다. Fig. 16은 AUC값이 0.5인 경우와 비교하여 민감도와 특이도가 확연히 더 높은 결과를 보여준다.

Table 7은 2017년 엔씨소프트사 게임의 데이터 마이닝 챌린지에 출전한 각 팀의 이탈 예측 모델의 정확도를 비교한 표이다[22]. 대표적인 MMORPG인 블라인드

앤 소울을 대상으로 하였다. 각 팀의 F-measure와 비교하여 본 모델은 평균 0.73의 더 높은 F-measure 값을 보여주었다.

5. 결론

가상 사회 내에서의 소셜 활동 및 사회적 관계는 사용자의 게임 참여도 및 흥미에 영향을 미칠 수 있으며 이탈 고객을 예측할 수 있는 기준이 될 수 있다. 본 논문에서는 클러스터링을 통해 길드 내 소속감 및 소셜 성향에 따라 사용자를 분류하고, 각 사용자 그룹을 대상으로 이탈 직전 다양한 게임 활동의 추세를 분석하여 소셜 활동과 이탈 원인과의 상관관계를 알아보았다. 또한 비슷한 소셜 활동 성향으로 구성된 각 그룹을 대상으로 이탈 직전 게임 활동에 대한 패턴을 추출하고, 이를 통해 최종적으로 이탈을 보다 정확하게 예측할 수 있는 모델을 제안하였다. 실험 결과, 대표적으로 개인플레이 길드의 Random Forest 알고리즘의 경우, TPR은 약 80%, FPR은 약 23%으로 기존 연구 결과보다 더 좋은 이탈 예측 성능을 보여주었다.

References

- [1] Global online gamingmarket, (2015, June 25). [Online]. Available: <http://www.marketwatch.com/story/global-online-gaming-market-2014-2014-06-25>
- [2] Williams, N. Yee, and S. E. Caplan, "Who plays,

- how much, and why? Debunking the stereotypical gamer profile," *Journal of Computer Mediated Communication*, Vol. 13, No. 4, pp.993-1018, Jul. 2008.
- [3] Blattberg, R.C. and J. Deighton, "Manage Marketing by the Customer Equity Test," *Harvard Business Review*, Vol. 74, No. 4, pp. 136-144, 1996.
- [4] Fleming, J.H., C. Coffman, and J.K. Harter, "Manage Your Human Sigma," *Harvard Business Review*, Vol. 83, No. 7, pp. 106-114, 2004.
- [5] Pritchard, J.O., "A Member's Lifetime Value," *Association Management*, Vol. 43, No. 6, pp. 35-39, 1991.
- [6] Crosby, L.A. and N. Stephens, "Effects of Relationship Marketing on Satisfaction, Retention, and Prices in the Life Insurance Industry," *Journal of Marketing Research*, Vol. 24, No. 4, pp. 404-411, 1987.
- [7] Lattin, J.M. and L. McAlister, "Using a Variety-seeking Model to Identify Substitute and Complementary Relationships Among Competing Products," *Journal of Marketing Research*, Vol. 22, No. 3, pp. 330-339, 1985.
- [8] Neslin, S., S. Gupta, W. Kamakura, J. Lu, and C. Mason, "Defection Detection: Measuring and Understanding the Predictive Accuracy of Customer Churn Models," *Journal of Marketing Research*, Vol. 43, No. 2, pp. 204-211, Aug. 2006.
- [9] Jaya Kawale, Aditya Pal, Jaideep Srivastava "Churn Prediction in MMORPGs: A Social Influence Based Approach," *Computational Science and Engineering, CSE '09. International Conference on*, Vol. 4, pp. 423-428, Aug. 2009.
- [10] Borbora, Zoheb, et al., "Churn prediction in mmorpgs using player motivation theories and an ensemble approach," *Privacy, Security, Risk and Trust (PAS-SAT)*, IEEE, pp. 157-164, 2011.
- [11] P. Datta, B. Masand, D. R. Mani, and B. Li, "Automated cellular modeling and prediction on a large scale," *Artif. Intell. Rev.*, Vol. 14, No. 6, pp. 485-502, 2000.
- [12] Yaya Xie Xiu Li E.W.T.Ngai, Weiyun Ying, "Customer churn prediction using improved balanced random forests," *Expert Systems with Applications*, Vol. 36, No. 3, pp. 5445-5449, Apr. 2009.
- [13] Chih-Fong Tsai, Yu-Hsin Lu, "Customer churn prediction by hybrid neural networks," Vol. 36, No. 10, pp. 12547-12553, Dec. 2009.
- [14] Bingquan Huang, Mohand TaharKechadi, Brian Buckley, "Customer churn prediction in telecommunications," Vol. 39, No. 1, pp. 1414-1425, Jan. 2012.
- [15] Yu Zhao, Bing Li, Xiu Li, Wenhuan Liu, and Shouju Ren, "Customer Churn Prediction Using Improved One-Class Support Vector Machine," *Advanced data mining and applications*, pp. 731-731, 2005.
- [16] Zoheb H Borbora, Jaideep Srivastava, "User Behavior Modelling Approach for Churn Prediction in Online Games," *ASE/IEEE International Conference on Privacy, Security, Risk and Trust*, pp. 51-60, 2012.
- [17] Julian Runge, Peng Gao, Florent Garcin, Boi Faltings, "Churn prediction for high-value players in casual social games," *Computational Intelligence and Games(CIG)*, pp. 1-8, Aug. 2014.
- [18] Kook-Yong Lee, "Factors Affecting Members' Sense of Belonging in Virtual Community," *Journal of the Korean Operations Research and Management Science Society*, Vol. 35, No. 1, pp. 19-45, 2010.
- [19] wikipedia, Bootstrap aggregating, (2017, August15). [Online]. Available: https://en.wikipedia.org/wiki/Bootstrap_aggregating
- [20] wikipedia, Random forest, (2017, October 11). [Online], Available: https://en.wikipedia.org/wiki/Random_forest
- [21] wikipedia, C4.5 algorithm, (2017, October 13). [Online], Available: https://en.wikipedia.org/wiki/C4.5_algorithm
- [22] Kim, Kyung-Joong, et al., "Game Data Mining Competition on Churn Prediction and Survival Analysis using Commercial Game Log Data," *arXiv preprint arXiv:1802.02301*, 2018.



서 은 비

2016년 숙명여자대학교 IT공학부 졸업(학사). 2018년 고려대학교 정보보호학과 졸업(석사). 관심분야는 데이터 마이닝, 딥러닝, 차량 보안, 게임 보안



우 지 영

2000년 KAIST 산업공학과 졸업(학사)
2002년 KAIST 산업공학과 졸업(석사)
2006년 KAIST 산업공학과 졸업(박사)
2006년~2008년 삼성전자 CRM파트 과장. 2008년~2010년 미국 아리조나대학 인공지능연구실 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 부교수. 2015년~현재 고려대학교 정보보호대학원 부교수. 관심분야는 온라인 게임 보안, 네트워크 보안, 네트워크 포렌식