ETRI

2017
Electronics and
Telecommunications
Trends

모바일 게임 분석 기술 동향

Mobile Game Analytics Technology Trends

Ⅰ. 서론
Ⅱ. 게임운영지표 분석 기술
Ⅲ. 게임행동예측 기술
Ⅳ. 결론

이상광 (S.K. Lee, sklee@etri.re.kr) 지식이러닝연구그룹 책임연구원
장시환 (S.H. Jang, jjangshan@etri.re.kr) 지식이러닝연구그룹 연구원
양성일 (S.I. Yang, siyang@etri.re.kr) 지식이러닝연구그룹 책임연구원

* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2017년도 문화기술 연구개발 지원사업으로 수행되었음[R2016030046, 지능형 라이브 서비스를 위한 게임 운영 시나리오 최적화 플랫폼 기술 개발].

최근 모바일 게임 산업에서는 게임 이용 중 게임 내 필요한 아이템을 별도로 구매하는 부분 유료화 비즈니스 모델이 지속해서 성장하고 있다. 부분 유료화 게임은 이용자 측면에서 접근성이 용이하며, 게임 제공자 측면에서는 구매를 유도하는 방법들이 고도화되었다. 인기게임의 경우 부분 유료화로 인한 수입이 지속적으로 증가하고 있다. 본고에서는 모바일 부분 유료화 게임에 대해 게임 운영을 최적화하기 위한 기술들을 살펴보고자 한다. 먼저, 게임 현황 파악 및 분석을 위한 대표적인 게임운영지표 분석 솔루션들을 요약하고, 게임운영지표를 개선하기 위한 게임 이용자 행동예측 기술들을 소개한다. 또한, 최근 연구되고 있는 모바일 게임 분석 기술의 한계점을 돌아보고 향후 연구 방향에 관해 기술한다.

본 저작물은 공공누리 제4유형
출처표시+상업적이용금지+변경금지 조건에 따라 이용할 수 있습니다.

© 2017 한국전자통신연구원

Ⅰ. 서론

최근 모바일 게임은 대부분 부분 유료화(F2P: Free-to-play) 형태로 출시되고 있다. F2P 게임은 이용자가 무료로 설치하여 게임을 실행할 수 있고, 개인의 선택에 따라 게임 재화, 유료 아이템 등을 앱 내 구매(IAPs: In-app purchases)를 통해 획득할 수 있다. 따라서, 게임 이용자의 구매 관련 행동 패턴을 분석하여 IAPs 활성화 방안을 마련하는 것이 F2P 게임 모델의 필수 요소이다.

또한, 모바일 게임에 대한 이용자 유입 비용(UAC: User acquisition cost)이 증가함에 따라, 게임 이용자들이 지속해서 게임에 접속할 수 있도록 유지하는 것 또한 중요 이슈가 되었다. Swrve 네트워크 내 30개 이상의 모바일 게임을 대상으로 한 Mutanen의 연구[1]에 따르면 신규 이용자 중 19%는 단 한 번 게임을 실행했으며, 66%는 첫 실행 후 24시간 내 게임을 그만두었다. 이러한 이용자의 게임 이탈은 이용자 잔존율(URR: User retention ratio) 및 충성도와 밀접한 관계가 있다[2]. 게임운영자는 URR의 감소를 막기 위하여 게임 이용자의 행동 패턴을 분석하고 게임 이탈이 예상되는 이용자를 예측하여 그에 대응하는 서비스를 제공할 필요가 있다. 이처럼, 모바일 게임 운영을 최적화하기 위해서는 게임 운영지표 분석이 선행되어야 하며, 이에 대응하는 게임 이용자의 행동패턴을 분석하여 대상 이용자를 분류하고 그에 대응하는 서비스를 적재적소에 제공하는 것이 필요하다.

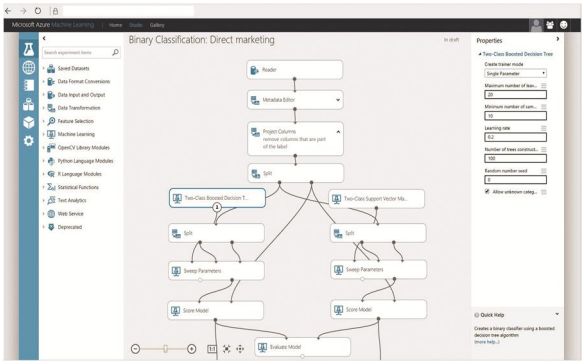
본고에서는 먼저, 게임운영지표 분석 기능을 제공하는 대표적인 솔루션들을 요약한 후, 게임운영지표 개선을 위한 모바일 게임 이용자 행동예측 기술의 최신 동향을 살펴보고자 한다.

Ⅱ. 게임운영지표 분석 기술

게임 운영지표 분석은 게이머들의 플레이로부터 발생되는 로그데이터를 수집하여, 게임 현황을 파악하고 게이머 행동 유형을 분류하는 기술이다. 이를 통하여 게임 현황을 시각화하고 관리자의 운영 편의성을 증대시킬 뿐만 아니라, 게이머 유형별 최적화 운영 시나리오를 제공하여 게임의 품질 및 게이머의 만족도를 향상하는데 활용한다.

1. Microsoft

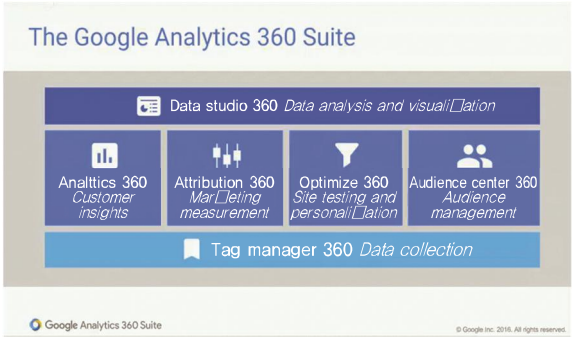
Microsoft는 엔터프라이즈급 클라우드 컴퓨팅 플랫폼 Azure 웹서비스를 개발하여 데이터베이스 관리 및 데이터 수집, 분석, 모니터링 등에 필요한 서비스와 API를 제공한다. 기계학습의 경우 기능별 순차블록을 조립하는 형태로 (그림 1)의 예시와 같은 GUI를 제공하여 사용자들의 편의성을 극대화 시켰으며, 최근 기계학습 기반 모바일 게임 데이터 예측 분석에 관한 기술강연, 웹 세미나 등을 통해 활발한 홍보 활동을 진행 중이다[3].



(그림 1) 기계학습 모델링 예시[3]

2. Google

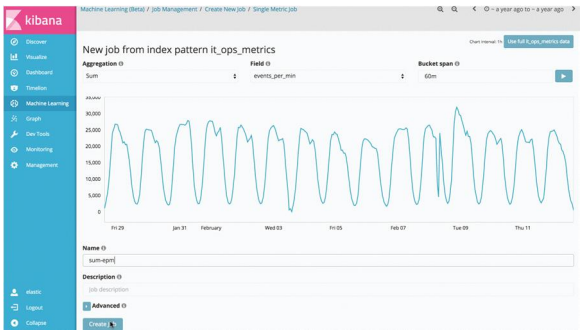
구글은 웹로그 분석 솔루션인 Google Analytics의 업그레이드 버전인 Google Analytics 360 Suite를 출시했다[(그림 2) 참조]. Google Analytics 360 Suite는 기업의 마케팅 성과를 통합 분석할 수 있는 기능이 생긴 것을 비롯해 보고서의 시각화 및 사용자 편의가 향상되었다[4].



(그림 2) Google Analytics 360 Suite 개요도[4]

3. Elastic

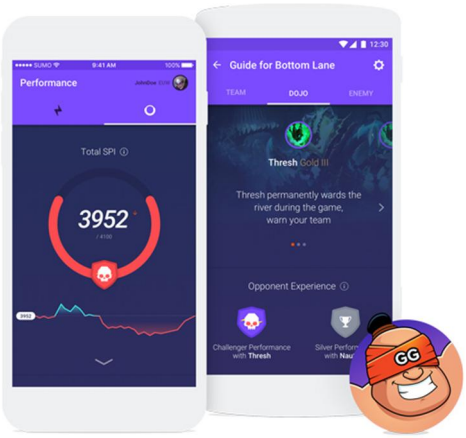
Elasticsearch는 유연한 플랫폼 기반으로 데이터 통계 분석에 장점이 있어 게임 분석을 위한 데이터베이스로 많이 활용되고 있다. Elastic은 최근 X-Pack을 통해 Elastic Stack의 기계학습 기능을 탑재하여 데이터로부터 통찰력을 얻을 수 있는 도구를 제공한다. (그림 3)과 같이, 현재의 X-Pack 기계학습 기능은 비지도(Unsupervised) 학습 방식을 통해 '시계열 데이터로부터 이상 징후 탐지' 기능을 제공하는 데 중점을 두고 있다[5].



(그림 3) 시계열에 따른 데이터의 이상 식별 예시[5]

4. DOJO Madness

e-스포츠 빅데이터 벤처기업인 DOJO Madness의 코칭 애플리케이션인 SUMO는 게임 데이터와 실시간으로 연동되어 게이머의 플레이를 분석하고 항목별 세부 정보 및 달성도의 통계치를 제공한다. 또한, 대전 상대

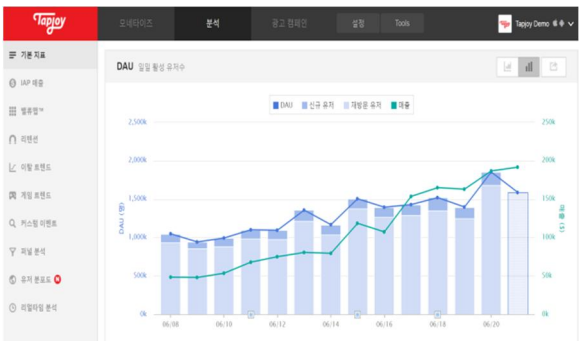


(그림 4) LoLSumo App 구동 예시[6]

의 성향 및 능력치, 직업조합, 아이템 등의 데이터를 종합적으로 분석하여 승리를 위한 맞춤형 공략법을 추천해준다. 대표적으로 (그림 4)와 같은 League of Legends(Riot Games)와 Overwatch(Blizzard)를 대상으로 서비스 중에 있다[6].

5. Tapjoy & 5Rocks

Tapjoy & 5Rocks는 분석과 마켓팅 운영 플랫폼을 사용하여 친구 수, 아이템 구매 액수, 유입 경로 등으로 실시간 코호트 분석(Cohort analysis)을 하고, 사용자 층을 세분화하여 타겟팅된 In-app promotion을 제공하는 등 게임의 수명연장과 수익창출 극대화를 위한 최적화 서비스를 제공한다[(그림 5) 참조]. 최근 오퍼월 플러스를 출시하여 모바일 광고 산업 육성에 주력하고 있다[7].



(그림 5) Tapjoy 분석 기본 지표 예시[7]

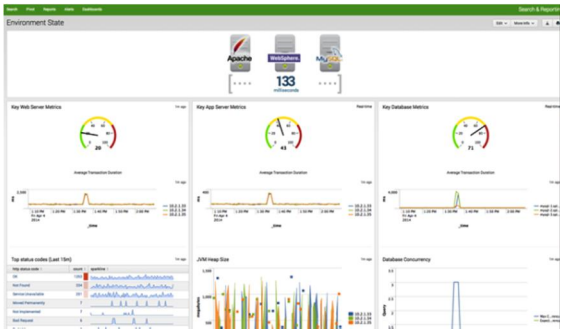
6. IGAWorks

IGAWorks는 계획수립, 운영, 통합 리포팅, 데이터 분석 기능이 포함된 솔루션을 통해 대상 콘텐츠의 리소스를 분석 및 관리하고 모바일비즈니스 및 마케팅 활동을 위한 가이드를 제공한다. 최근 모바일광고 업계 지도를 (그림 6)과 같이 공개하는 등 간접수익 창출을 위한 ADTech (Advertisement technology) 기업으로서의 자리를 매김하고 있다[8].

(그림 6) S.Korea Mobile Ad Technology Landscape[8]

7. Splunk

Splunk는 운영 인텔리전스를 위한 플랫폼으로 Splunk Enterprise를 제공하여 Machine-Generated Data에 기반을 둔 IT 시스템과 기술 인프라 현황을 실시간으로 확인하여 데이터를 가치 있게 사용할 수 있도록 지원한다[(그림 7) 참조], [9].



(그림 7) Splunk Enterprise의 모니터링 예시[9]

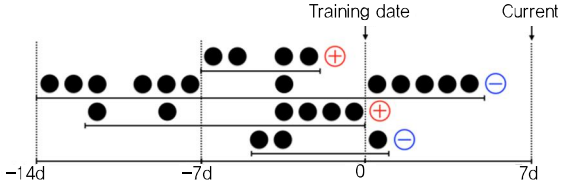
Ⅲ. 게임행동예측 기술

게임 이용자 행동 예측은 접속, 게임플레이, 구매 등 게임 이용자의 게임 내 활동이 기록된 게임로그데이터로부터 특정 행동에 관련된 특징을 추출하여 행동패턴을 분석하고 해당 행동을 예측하는 기술이다. 모바일 F2P 게임의 매출은 주로 광고 및 IAPs에 의해 발생된다. 따라서 게임 이용자의 행동을 모니터링하고 분석하며 예측하는 기술이 이 비즈니스모델의 중요한 열쇠가 되었다. 최근 모바일 F2P 게임에 대한 게임 이용자 행동 예측 기술이 활발히 연구되고 있으며[10]-[15], 본 장에서는 게임 이용자의 게임 이탈 및 구매 행동에 관한 예측기술을 살펴본다.

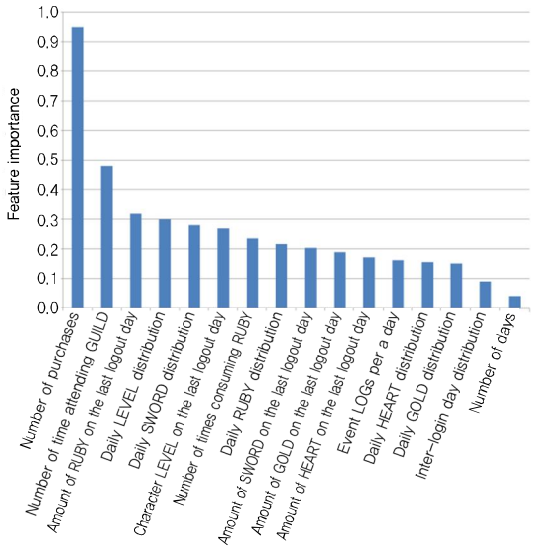
1. 게임 이탈 예측

모바일 게임이 성공하기 위한 가장 중요한 요소 중 하나는 게임을 떠날 게임 이용자를 분류하고 정의하는 역량이다. 이러한 형태의 분석을 게임 이탈 예측(Churn prediction)이라고 한다. 게임 서비스를 해지하는 이용자를 게임 이탈자(Churner)라고 하며, 게임 잔존자 대비 게임 이탈자의 비율을 게임 이탈률(Churn rate)이라고 한다. 일반적으로 F2P 게임에서 게임 이탈률은 매우 높으며, 대부분의 게임 이용자들이 첫 게임 플레이 시 게임을 이탈한다. 따라서 게임 이탈자를 예측하고 그들이 게임을 떠나지 않도록 장려하는 것이 필요하다.

Hadiji와 동료 연구자들[10]은 처음으로 F2P 게임에 대해 게임 이탈 예측을 시도하였다. 먼저, (그림 8)과 같이 관찰 시점 이전 7일 내 접속이 없었다면 게임 이탈로 정의하였다. (그림 8)과 (그림 10)에서 ●은 게임 활동



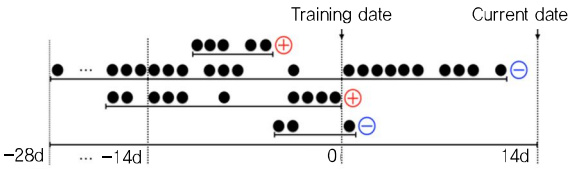
(그림 8) Hadiji와 동료연구자에 의한 게임 이탈 정의



(그림 9) 특징 간 상대적 중요도

날짜이며, ⊕는 게임 이탈, ⊖는 게임 잔존을 나타낸다. 일반적으로 모바일 게임의 경우 한 주를 주기로 게임운영지표 특성이 나타나기 때문에 게임 이탈 시점을 주 단위로 결정하는 것이 합리적이다. 또한, 게임 이용자 수 및 접속 빈도 등에 따라 게임 이탈 시점을 유연하게 적용할 필요가 있다. 게임 이탈 예측 기술의 가장 중요한 요소 중 하나는 게임 이용자의 게임 이탈을 예측할 수 있는 적절한 행동 특징(Feature)을 선택하는 것이다. Hadiji와 동료의 연구[10]는 게임 이탈에 영향을 주는 접속, 플레이 시간, 결제 등의 특징을 추출하였으며 결정트리(DT: Decision tree), 랜덤 포레스트(RF: Random forest), 신경망(NN: Neural network) 등의 분류기를 이용하여 게임 이탈 예측 모델링을 위한 학습을 수행하였다. 특히, RF의 경우 특징 간의 상대적 중요도를 제공해주며 이는 게임 이탈의 원인 분석을 위한 주요 단서를 제공해준다. (그림 9)는 엠게임의 크레이지드래곤 게임에 대한 특징 간 상대적 중요도 예를 보여준다.

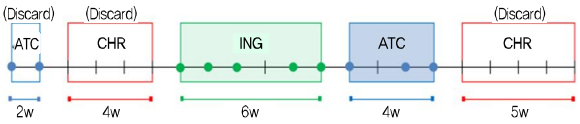
Runge와 동료들[11]은 모바일 F2P 게임에 대해 고가치 이용자를 정의하고 이들에 대한 게임 이탈 예측을 수행하였다. 이 연구에서는, 캐쥬얼 소셜 게임인 Wooga사의 Diamond Dash 게임에 대해 상위 7% 결제 이용자



(그림 10) Runge와 동료연구자에 의한 게임 이탈 정의

가 총 결제금액의 약 50%를 차지한다는 사실을 기반으로, 관찰시점 이전 90일간 상위 10% 결제 이용자를 고가치 이용자로 정의하였다. 또한, (그림 10)과 같이, 관찰시점 이전 14일 이상 게임에 접속하지 않을 경우 게임 이탈자로 간주하였다. Runge와 동료들[11]은 게임 이탈에 영향을 주는 행동 특징을 추출하였으며, DT, NN, SVM(Support vector machine), 로지스틱 회귀(LR: Logistic regression) 분류기를 이용하여 모델링하였다. 또한, 시간 특성을 반영하기 위해 은닉마르코프모델(HMM: Hidden Markov model)도 도입하였다. 성능 평가를 위해 세 개의 그룹으로 나누어 A/B 테스트를 수행하였으며, 고가치 이용자의 게임 이탈을 예측하여 인센티브를 제공하여도 게임 이탈률에는 큰 영향을 주지 않는다는 사실을 보여주었다.

Xie와 동료들[12]은 게임 이벤트 빈도수 기반의 게임 이탈 모델을 제시하였다. 제안 방식은 모든 게임에 존재하는 게임 이벤트에 대한 빈도수만을 이용하므로, 특정 게임에 종속되지 않고 좀 더 포괄적으로 적용될 수 있다. 이 연구에서는 게임 이탈을 Churn이라는 용어 대신 Disengagement라는 용어로 대체하여 사용하였다. 먼저, 게임 이탈(Disengagement)을 정의하기 위해 각 게임 이용자에 대해 지지난 한 달 및 지난 한 달 동안의 게임 내 총 활동량을 각각 계산하고 정렬한다. 총 활동량은 모든 이벤트 빈도수의 총합이다. 각 달에 대해 정렬된 총 활동량 리스트는 사분위수(Quartile) 상에 나뉘고 각 이용자는 그들의 활동량에 따라 1과 4 사이의 순위가 결정된다. 지지난 한 달 보다 지난 한 달의 순위가 2보다 크면 게임 이탈 그룹에 할당하고, 그렇지 않으면



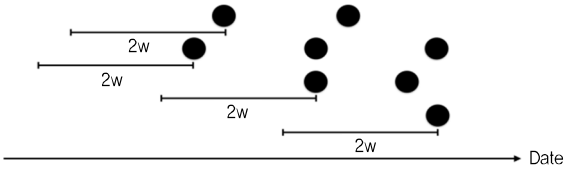
(그림 11) Tamassia와 동료들에 의한 시계열 행동 데이터 시퀀스

미이탈 그룹에 할당한다. 즉, 한 달 단위로 활동량을 계산하여 활동량이 일정 수준 낮아지면 게임 이탈로 간주하게 된다. Xie와 동료들[12]은, 게임 이탈 모델링을 위해 DT, LR, SVM을 이용하였으며, Runge와 동료들[11]의 방식보다 성능이 우수함을 보여주었다.

Tamassia와 동료들[13]은 게임 이탈을 게임 이용자의 행동 특성에 대한 시계열 분석으로 예측하였다. 이 연구에서는 플레이 시간과 평균 미접속 일수의 분포를 분석하여 게임 이탈 시점을 4주로 설정하다. 그리고 각 게임 이용자에 대해 시간의 역방향 흐름에 따라 4주 내 플레이하지 않았을 경우 CHR(Churned), CHR 이전 4주 내 접속했다면 ATC(About-to-Churn), 그 외 ING(In-Game)로 레이블링하였다[(그림 11) 참조]. 이때, CHR과 4주 미만의 ATC는 모두 버리고 ATC를 양성 샘플(Positive sample), ING를 음성으로 하였으며 이렇게 구성된 시퀀스는 HMM 분류기를 이용하여 학습하였다. 제안 방식의 성능 평가를 위해 DT, RF, LR 등의 비순차적 분류기(Non-temporal classifier)와 비교하였으며 대부분의 비순차적 분류기에 비해 우수한 AUC(Area under the curve)를 나타내었다.

2. 구매 예측

일반적으로 F2P 모바일 게임에서는 유료 서비스를 결제하지 않는 무과금 이용자가 대다수를 차지하고 있다. 따라서 F2P 게임의 주요 도전과제는 무과금 이용자를 결제 이용자로 전환시켜, 이용자를 가입시키는데 드는 비용인 UAC보다 가입된 이용자의 가치인 고객평생가치(LTV: Lifetime value)가 더 높도록 하는 것이다.



(그림 12) Xie와 동료들에 의한 구매 정의

Xie와 동료들[12]은 게임 이용자의 과거 2주간 행동을 관찰하여 이용자의 첫구매를 예측하였다. (그림 12)에서 ●은 구매 날짜를 나타낸다. 즉, 첫 구매 전 2주간의 행동 데이터를 결제 이용자에 대한 학습에 이용하였으며, 구매가 없는 이용자에 대해서는 임의로 선택된 2주간의 행동 데이터를 통해 학습하였다. 2주라는 단위 기간은 이용할 수 있는 게임로그 데이터양에 따라 결정하였다. 실제로, 이용 가능한 데이터양에 따라 기간을 조정할 수 있다. Ⅲ장 1절에서 언급했던 게임 이탈 예측과 동일하게, [12]에서는 게임 이벤트의 빈도수만을 이용하여 결제 여부를 예측하기 때문에 모든 게임에 대해 좀 더 포괄적으로 적용될 수 있다. 성능 평가를 위해, DT, LR, SVM 분류기를 이용하여 모델링하였으며, 기존의 방식에 비해 우수한 AUCPR(Area under the precision-recall curve)을 나타내었다. 하지만, 게임 이용자의 전체 로그데이터가 아닌 이용 가능한 데이터 내에서 처음 발생된 구매를 첫 구매로 가정하였기 때문에 실제 첫 구매 예측과는 다소 상이하다는 한계가 있다.

Sifa와 동료들[14]은 게임 이용자의 구매여부뿐만 아니라 결제금액에 관한 예측도 동시에 수행하였다. 먼저 구매관련 행동 특징을 추출한 후, DT, RF, SVM 분류기를 통해 모델링하여 구매여부 예측을 수행하였으며, 포아송회귀트리(Poisson regression tree)를 통해 구매량 예측 모델링을 하였다. 위의 모델링을 통해, 신규 이용자가 3일 내 유료 아이템을 결제하면 향후 평균 이상 추가 구매할 가능성이 높다라는 사실을 보여주었다.

3. 게임 데이터 불균형 문제

모바일 F2P 게임의 대다수가 게임 이탈 확률이 높거

〈표 1〉 게임 이탈자 및 게임 잔존자 분포

데이터세트	크레이지드래곤
게임 이용자 수	5,710
게임 이탈자 수	3,687
게임 잔존자 수	2,023

나 무과금 이용자이기 때문에, Ⅲ장 1절과 Ⅲ장 2절에서 기술한 게임 이탈 및 구매 예측을 모델링을 하는 데 있어 심각한 데이터 불균형 문제가 존재한다. 즉, 분류기를 통한 학습 과정에서 상대적으로 구성원이 적은 소수의 클래스에 대해 모델링이 제대로 되지 않는다. 이러한 데이터 불균형 문제를 해결하기 위해 [11]에서는 게임 이탈 예측 모델링에 대해 Box-Cox 변환 방식을 이용하였다. 마찬가지로, [14]에서는 구매 예측 모델링에 대해 결제 이용자 및 무과금 이용자의 불균형 분포를 해결하기 위해 SMOTE(Synthetic minority over-sampling technique)[16] 방식을 적용하였다. SMOTE는 소수 클래스들 집합 사이에 새로운 데이터를 추가하여 오버샘플링(Oversampling)하는 방법으로, [15]에서도 게임 이탈자에 비해 상대적으로 적은 게임 잔존자 수를 극복하기 위해 이 방식을 이용하였다. 〈표 1〉은 엠게임의 크레이지드래곤 게임에 대한 게임 이탈자 및 게임 잔존자 구성을 보여준다. 〈표 1〉에 보여주듯이, 특정 기간에 대한 게임 잔존율은 약 35.4%로서, 게임 이탈률과 균형을 이루지 못하고 있다.

Ⅳ. 결론

본고에서는 모바일 게임 분석을 위한 게임운영지표 분석 기술과 게임행동예측 기술의 동향을 살펴보았다. 모바일 F2P 게임은 게임 이용자의 잔존과 구매 활성화 등에 대한 게임운영 최적화가 요구되는 비즈니스 모델로서, 게임운영지표 분석을 통해 게임 현황을 파악하고, 서비스 개선을 위해 게임 이용자의 행동을 분석하고 예측하는 기술이 필수이다. 상기 기술한 내용과 같이, 게임운영지표 분석 기술은 주로 상용화 솔루션을 통해 다양하게 제공되고 있지만, 게임 이용자의 행동예측 기술은 게임 이탈 및 구매 예측에 관한 연구에만 국한되어 확장되지 못하고 있다. 향후 이러한 게임 행동예측 기술들을 게임 라이브 서비스에 실제 적용하기 위한 다각적인 연구가 진행되어야 하며, 연구 활성화를 위해 양질의 게임로그데이터 공개 방안도 동시에 마련되어야 한다.

용어해설

부분 유료화 게임(Free-to-play game) 게임 이용자가 무료로 설치하여 게임을 플레이할 수 있으며, 선택적으로 게임 재화, 유료 아이템 등을 앱 내 구매(IAPs: In-app purchases,)를 통해 획득할 수 있는 게임 비즈니스 모델

약어 정리

ADTech	Advertisement Technology
AUC	Area under the Curve
AUCPR	Area under the Precision-Recall Curve
DT	Decision Tree
F2P	Free-to-Play
HMM	Hidden Markov Model
IAPs	In-App Purchases
LR	Logistic Regression
LTV	Lifetime Value
NN	Neural Network
RF	Random Forest
SMOTE	Synthetic Minority Over-Sampling Technique
SVM	Support Vector Machine
UAC	User Acquisition Cost
URR	User Retention

참고문헌

[1] SWRVE, "The SWRVE New Players Report," Accessed 2017. http://www.swrve.com/whitepapers/files/SWRVE-new players-report-2015.pdf

[2] T. Mutanen, J. Ahola, and S. Nousiainen, "Customer Churn Prediction – a Case Study in Retail Banking," in *Proc. ECML/PKDD 2006 Workshop on Practical Data Mining: Applications, Experiences and Challenges*, Sept. 2006, pp.

13–19.

[3] Microsoft Azure, Accessed 2017. https://azure.microsoft.com

[4] Google, Accessed 2017. https://www.google.com

[5] Elastic, Accessed 2017. https://www.elastic.co

[6] DOJO MADNESS, Accessed 2017. http://dojomadness.com

[7] 5Rocks, Accessed 2017. http://www.5rocks.io

[8] Igaworks, Accessed 2017. http://www.igaworks.com

[9] Splunk, Accessed 2017. https://www.splunk.com

[10] F. Hadiji et al., "Predicting Player Churn in the Wild," in *Proc. IEEE Conf. Comput. Intell. Games*, Dortmund, Germany, Aug. 26–29, 2014, pp. 1–8.

[11] J. Runge et al., "Churn Prediction for High-Value Players in Casual Social Games," in *Proc. IEEE Conf. Comput. Intell. Games*, Dortmund, Germany, Aug. 26–29, 2014, pp. 1–8.

[12] H. Xie et al., "Predicting Player Disengagement and First Purchase with Event-Frequency Based Data Representation," in *Proc. IEEE Conf. Comput. Intell. Games*, Tainan, Taiwan, Aug. 31–Sept. 2, 2015, pp. 230–237.

[13] M. Tamassia et al., "Predicting Player Churn in Destiny: a Hidden Markov Models Approach to Predicting Player Departure in a Major Online Game," in *Proc. IEEE Conf. Comput. Intell. Games*, Santorini, Greece, Sept. 20–23, pp. 325–332.

[14] R. Sifa et al., "Predicting Purchase Decisions in Mobile Free-to-Play Games," in *Proc.Eleventh AAAI Conf. Artif. Intell. Interactive Digital Entertainment (AIIDE 2015)*, Santa Cruz, CA, USA, Nov. 14–18, 2015, pp. 79–85.

[15] S. Lee et al., "Predicting Churn in Mobile Free-to-Play Games," in *Proc. Int. Conf. Inform. Commun. Technol. Convergence*, Jeju, Rep. of Korea, Oct. 19–21, 2016, pp. 1046–1048.

[16] N.V. Chawla et al., "SMOTE: Synthetic Minority Over-Sampling Technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, Jan. 2002, pp. 321–357.