

리뷰-피드백 프로세스를 통한 고객 이탈률 추정: 텍스트 마이닝, 계량경제학, 준실험설계 방법론을 활용한 실증적 연구

Estimate Customer Churn Rate with the Review-Feedback Process: Empirical Study with Text Mining, Econometrics, and Quai-Experiment Methodologies

김 초 이 (Choi Kim)	경희대학교 경영학과 석사과정
김 재 민 (Jaemin Kim)	경희대학교 빅데이터경영학과 석사과정
정 가 현 (Gahyung Jeong)	경희대학교 빅데이터경영학과 석사과정
박 재 흥 (Jaehong Park)	경희대학교 경영대학 교수, 교신저자

요 약

기존 연구들은 주로 사용자의 게임 참여 동기나 사회적 욕구에 따른 이탈 요인을 연구하였다. 하지만, 기존 연구들은 게임 참여 동기 관점에서 집중하다 보니, 사용자 불만 사항 개선에 따른 사용자 이탈에 관한 분석은 비교적 적게 이루어져왔다. 게임에 대한 사용자 불만 사항과 그에 따른 게임 품질 개선은 사용자가 게임에 참여하는 요인 중 하나이다. 따라서, 본 연구는 사용자 불만 요인이 사용자 이탈에 미치는 영향을 실증적으로 분석하여 그 관계를 살펴보고자 한다. 본 연구는 최근 유행했던 “PUBG - 배틀그라운드 게임”을 분석하여 제품 품질에 대한 불만 사항 피드백이 얼마나 사용자 이탈에 영향을 주는지 실증적으로 분석한다. 텍스트 마이닝(Text Mining) 분석을 통해, 사용자들의 품질에 대한 불만요인을 도출하였고, 콕스모델(Cox Model)을 통해 불만 요인에 따른 사용자의 이탈률을 추정하였다. 또한 준실험설계 방법을 통해 실제 불만사항 개선 패치에 따라 사용자 수가 어떻게 변화하는지 살펴봄으로 본 연구 결과를 검증하였다. 분석 결과, 불만 사항 중 게임의 재미와 관련된 요인들이 사용자 이탈에 가장 큰 영향을 주었고, 반면 게임 사용 편의성과 관련된 불만 사항들은 비교적 사용자 이탈에 적은 영향을 준다는 것을 실증적으로 보였다. 본 연구결과에 따르면, 게임 불만 요인 개선에 따라 사용자들의 이탈 정도가 달라질 수 있으며, 이에 따라 게임 사용자 관리를 할 수 있다는 것을 밝혀냈다. 본 연구는 게임 개발 및 운영사 입장에서 사용자들의 불만 사항 개선에 대한 우선 순위를 제공해 줌으로서 실증적인 공헌을 제시한다.

키워드 : 리뷰-피드백 프로세스, 사용자 이탈률 분석, 게임 수명 예측, 텍스트 마이닝, 콕스 비례-위험 모델, 시계열 분석, 준실험설계, 비교 투입 시계열 설계

† 이 논문 또는 저서는 2019년 대한민국 교육부와 한국연구재단의 인문사회분야 중견연구자지원사업의 지원을 받아 수행된 연구임(NRF-2019S1A5A2A01046660).

I. 서 론

최근 국내 게임 업체는 시장 규모를 확대하고 있다. 2018년 기준 국내 게임 시장의 규모는 전년 13조 1,423억 대비 8.7% 증가한 14조 2,902억 원으로 꾸준한 성장세를 보였다. 부가가치 측면에서 보면 2018년 기준 부가가치액은 6조 1,791억 원, 부가가치율은 43.2%로 나타났다(한국콘텐츠진흥원, 2018, 2019). 이처럼 게임 산업은 지속적인 성장세를 보이고 있으며 동시에 높은 부가가치를 창출하고 있다. 이에 따라 게임 산업은 전체 콘텐츠 산업의 성장을 견인하는 핵심 산업으로 주목받고 있다.

게임산업의 제품인 게임은 제품개발비용이 많이 요구되는 반면, 출시 이후 지출되는 유지보수 비용은 상대적으로 적다는 특징을 갖는다. 이 때문에 게임산업은 대표적인 고위험-고부가 가치 산업으로 여겨진다. 개발이 완료된 후에는 이용자인 명에게 추가로 서비스를 제공하기 위한 한계비용이 거의 제로에 가깝다. 그렇기 때문에 개발사들은 출시된 게임이 어느 정도 흥행에 성공했다면 별다른 투자 없이도 지속적으로 수익을 창출할 수 있다(삼정KPMG, 2018). 이는 개발사들이 시장에서 살아남기 위해서는 신규 사용자를 유입시키거나 기존 사용자를 유지하는 것에 주력해야 한다는 시사점을 준다. 다시 말해 장기적인 관점에서 사용자의 불만 사항들을 관리하는 것이 제품의 성공에 중요한 요인이라는 의미다.

게임이 장기적으로 시장에서 살아남을 수 있는가 여부는 그 게임이 얼마만큼 지속적으로 사용자를 보유하고가에 달려있다(Cheung *et al.*, 2015; Evans, 2016). 게임산업의 경우 주로 게임판매와 게임 내 콘텐츠 판매를 통해 수익이 창출된다. 따라서 게임이 출시된 초기 이후의 수익 대부분은 사용자의 게임 내 결제를 통해 창출된다. 이러한 이유로 게임 업체들은 사용자의 이탈 방지를 통해 경쟁 우위 달성하고자 한다. 사용자 이탈에 대한 기존의 연구들은 사용자의 게임 참여 동기(motivation)를 통해 사용자의 이탈을 예측하고자 하였

다. 주요 변수로는 사용자의 개인적 동기(성취, 현실 도피 등), 그리고 게임 내 사용자의 사회적 욕구(파티, 길드 등)와 같은 사용자 개별적 특성과 사용자의 인구통계학적 요소들이 주로 활용되었다(Griffiths *et al.*, 2004; Yee, 2006).

그러나, 게임의 품질 변화에 따른 사용자의 이탈은 방법론적인 문제로 상대적으로 연구가 활발하지 않은 분야였다. 하지만, 사용자들의 요구사항과 불만 요인들을 적절히 반영하여 게임의 품질을 유지하는 것은 사용자의 이탈과 직접적인 연관이 있다는 측면에서 연구가 필요한 분야이다. 일례로 국내에서 출시된 온라인 게임에서 게임 아이템 및 거래 시스템 등에 대한 사용자 불만이 폭발하여 대규모의 사용자들이 이탈한 사례가 있었다. 해당 온라인 게임은 상당한 인기와 매출을 자랑하던 게임이었지만, 사용자 이탈이 가속화됨에 따라 해당 게임사는 국내 시장에서 기업 운영에 치명적인 위기를 맞이하게 되었다. 이는 게임 업계에 고객 불만 사항 관리의 중요성을 일깨워준 대표적인 사례 중 하나이다(게임메카, 2011). 하지만, 사용자의 게임에 대한 불만 관리의 중요성에도 불구하고, 과거에는 사용자 불만 요인에 대한 변수들의 실제 데이터 수집이 쉽지 않았으며 분석 방법론의 활용이 어려웠기 때문에 게임 내 사용자 불만 요인과 관련된 많은 연구가 진행되지 못하고 있었다.

이에 본 연구는 게임의 품질에 대한 사용자들의 불만 요인이 사용자 이탈에 어떤 영향을 주는지 밝히고자 한다. 먼저, 사용자수(Peak Concurrent Connected User, 최고 동시 접속자 수) 데이터, 개별 사용자의 플레이 시간 데이터, 그리고 리뷰 텍스트를 일단위로 웹 크롤링을 통해 수집하였다. 리뷰 텍스트의 경우 text mining 기법을 활용하여 불만 요인 키워드를 추출하였다. 그리고 추출된 요인들을 cox hazard model의 예측변수로 활용하였다. 이를 통해 각 불만 요인들의 이탈위험 증가율을 계산하였다. 더 나아가 준실험설계(quasi-experiment) 분석 방법을 통해, 실제 사용자 불만 사항이 반영된 패치를 통해 게임 품질이 개선되었을

때 실제 사용자 이탈이 어느 정도로 방지되는지 살펴봄으로써 본 연구 결과를 검증하였다.

본 연구는 배틀그라운드 게임을 연구 대상으로 선정하고, 해당 게임의 데이터를 약 1년간 수집하여 총 35,991건의 사용자 행태 데이터를 구성하였다. 리뷰를 통해 추출된 불만 요인 중 게임 재미와 관련된 요인들이 사용자 이탈에 가장 큰 영향을 주는 것을 발견하였다. 예를 들어, 게임의 규칙 및 설계가 지루하게 설정되었다는 불만 요인은 하루에 약 22.4%씩 사용자 이탈 위험을 높일 수 있다. 반면 게임 사용 편의성과 관련된 UX/UI 불만 요인은 약 5.3%의 일평균 이탈 위험도 증가에 기여함에 따라 비교적 적은 영향을 주는 것을 보였다. 이런 결과들을 검증하기 위해 준실험방법론을 활용하여 실제 패치 내용을 살펴본 결과, 불만요소를 개선하는 패치가 실제로 사용자 수 감소 추세를 완화하고 게임 사용자 수에 긍정적인 영향을 미치는 것을 확인하였다.

이에 따라, 본 연구는 text mining과 cox hazard model을 결합하여, 사용자의 불만요인에 따른 사용자이탈을 예측하는 새로운 분석 방법론을 제시하였다. 리뷰데이터를 사용한 기존의 연구들은 리뷰데이터를 활용하여 리뷰가 갖는 구전효과(word-of-mouth)를 관찰하였다. 리뷰데이터를 감정어 사전에 따라 긍정, 부정으로 분류하고, 긍정리뷰, 부정리뷰, eWOM(e-Word of Mouth, 온라인 구전효과)등을 독립변수로 두어 이에 따른 판매량을 연구하였다(Jung *et al.*, 2014). 즉, 리뷰 데이터가 다른 사용자의 제품 및 서비스 구매 결정에 미치는 영향을 분석한 것이다. 하지만 리뷰데이터를 통해 표출되는 사용자의 제품에 대한 불만사항이 제품 서비스 품질개선 절차에 반영되었을 때, 사용자의 이탈에 대한 영향에 대한 연구는 이루어지지 않았다. 따라서, 본 연구는 텍스트 마이닝 기법을 통해, 리뷰데이터에서 서비스 품질 키워드를 도출해 우선순위에 따라 분류하고, 불만 키워드의 내용에 대한 패치가 이루어졌을 때, 해당 사용자 이탈에 미치는 영향을 생존분석을 통해 살펴보고

자 한다. 또한, quasi-experiment 방법론을 활용하여, 사용자 불만 사항 개선에 따른 이탈률을 비교하여 본 연구 방법론의 타당성을 검증하였다.

이에 본 연구는 리뷰를 사용자의 불만 사항을 도출하는데 활용하였고, 불만 사항 중 게임의 재미와 관련된 요인들이 사용자 이탈에 가장 큰 영향을 준다는 것을 실증적으로 밝혔다. 또한 게임 사용 편의성과 관련된 불만 사항들은 비교적 사용자 이탈에 적은 영향을 준다는 것을 보였다. 본 연구는 게임 개발 및 운영사 입장에서 사용자들의 불만 사항 개선에 대한 우선 순위를 제공해 줌으로써 실증적인 공헌을 제시할 수 있다.

본 논문의 구성은 다음과 같다. 제II장에서는 선행연구의 흐름을 파악하고, 게임 품질이 사용자 이탈에 미치는 영향에 대한 가능성을 제시한다. 제III장에서는 연구 데이터와 데이터의 특성에 대해 설명한다. 제IV장에서는 본 연구에 이용된 연구 방법론(텍스트 마이닝(Text-Mining), 콕스 모델링(Cox-Modeling), 준실험설계(Quasi Experimental Design))에 대해 기술하였다. 제V장 이후로는 연구의 결과와 의의 한계점에 대해 제시한다.

II. 문헌연구

게임 사용자의 수는 게임사들의 수익 창출에 직접적으로 기여하므로, 게임사들은 사용자들이 자사의 서비스를 지속적으로 이용하도록 사용자 이탈을 방지하기 위해 노력해왔다. 따라서 게임 내 사용자의 변동을 추적하는 사용자 이탈 예측분석(churn prediction analysis)은 게임 산업분야에서 주요 연구 주제가 되어 왔다(Borbora and Srivastava, 2012; Hadiji *et al.*, 2014; Kawale *et al.*, 2009). 기존 사용자이탈을 예측하는 분석 연구에서는 이탈의 주요 요인 중 하나로 사용자의 동기를 연구해 왔다. 사용자의 게임 참여 동기는 게임 내 사용자의 전반적인 행동 양상을 예측하는 중요한 요인으로 작용한다. 즉 게임 참여 동기를 예측할 수 있다면, 사용자의 행동 변화를 사전에 탐지하여 이탈을 방지할

수 있음이 선행연구를 통해 연구된 바 있다(Kawale *et al.*, 2009).

또한 사용자의 게임 참여 동기 분석 외에 다른 방법으로 사용자의 이탈을 예측 또는 방지하는 연구가 이루어지기도 했다. 이를 위한 방법론으로는 RFM(Recency, Frequency, Monetary Value)과 TFPD 분석(time-frequency plane domain analysis)이 쓰였다. RFM은 간단한 시계열적 특성을 가져 온라인 게임 사용자의 로그인 기록만을 기반으로 합리적인 예측 모형을 제안 가능하다. 더불어 TFPD 분석의 경우 RFM에 비해 20% 이상의 성능을 보일 수 있음이 연구되었다(Castro and Tsuzuki, 2015). 사용자의 이탈을 방지하는 방법론 또한 연구되었는데, 사용자활동 분석을 통한 개인화 푸시 알림으로 사용자이탈을 약 28% 감소시킬 수 있다는 것을 보였다(Milošević *et al.*, 2017).

최근 게임 참여 동기에 관한 기존 연구는 크게 두 방향으로 나뉘고 있다. 첫 번째 연구 흐름은 게임적 보상 및 요구 사항 개선을 통한 사용자의 게임에 대한 애착 및 충성도 형성이 게임 플레이에 대한 동기에 미치는 영향에 대한 것이다. 먼저 Wu *et al.*(2010)은 게임 플레이에 대한 보상(gratification)과 시스템 설계가 사용자의 게임 플레이를 지속하려는 동기에 대해 긍정적인 영향을 미치는 것을 보였다. 또한 Jung *et al.*(2014)은 게임의 기술적인 혁신성과 사용자 중심의 디자인이 게임 플레이 의향을 높이는 것을 보여주었다. 그리고, Moon *et al.*(2013)에 따르면 사용자의 게임 충성도를 높이는데 효과적인 전략은 1) 게임 캐릭터에 대한 권한을 높이고, 2) 가상의 온라인 신분을 만들어 주는 것이라 밝혔다.

두 번째 연구 흐름은 게임 사용자간 상호작용을 통해 발생하는 사회적 욕구가 게임 플레이 및 참여에 미치는 영향에 대한 것이다. Chen *et al.*(2006a, 2006b)은 게임 내 사용자 간 사회적 교류가 게임을 지속적으로 플레이하게 하는 주요 요인임을 밝혔다. 또한 Martončík and Lokš(2016)은 MMORPG(Massively Multiplayer Online Role-Playing Game)

내의 길드 활동을 통해 사용자들이 외로움과 불안 장애를 완화할 수 있음을 보이기도 했다. 사용자 이탈과 관련한 연구들에서도 게임 플레이와 사회적 요인을 고려하여 사용자의 게임 플레이에 대한 동기를 분석하고 이를 사용자 이탈과 관련해 분석하는 등의 다양한 시도가 있었다(Borbora and Srivastava, 2012; Kawale *et al.*, 2009).

하지만, 이 두 가지 흐름의 연구들은 모두 사용자 불만 사항들에 대한 게임 서비스 개선이 사용자 이탈에 미치는 영향을 고려하지 못한다는 한계점이 있다. 앞서 말했던 Jung *et al.*(2014)의 연구에서, 게임을 지속적으로 플레이하는 사용자들은 사용자의 불만사항(혹은 요구사항)을 깊이 이해하여 설계된 게임의 기능들과 상호작용하기 원한다고 보고했다. 즉, 사용자 의견이 반영된 패치 업데이트는 사용자의 게임 참여 동기를 크게 개선할 수 있을 것이다. 이는 지속적으로 사용자들을 확보하기 위해서는, 게임에 대한 사용자들의 불만사항 파악이 필수적이라는 것이다. 이에 본 연구에서는 게임 품질에 관련된 사용자 리뷰를 분석하여 중요 불만 요인들을 도출하고, 이 요인에 따른 사용자 이탈률을 추정하고자 한다. 또한, 이에 대한 개발사의 패치 및 게임 업데이트를 통한 개선이 사용자 이탈에 미치는 영향에 대해 연구함으로써 선행연구의 공백을 보완하고자 한다.

게임은 사용자 리뷰와 개발자의 피드백(업데이트)이 매우 활발하게(많은 경우 실시간으로) 이루어지는 변동성이 큰 소프트웨어 제품이다. 따라서 본 연구는 기존 연구 흐름들과 더불어 게임 서비스에 대한 사용자의 불만요인과 그에 대한 개발사의 피드백 효과를 고려함으로써 이전의 연구들을 보완하고 정확한 게임 사용자 이탈을 예측하고자 한다. 이를 위해 게임 불만 요인들을 게임 사용자 리뷰로부터 추출하고, 이를 활용해 특정 요인들이 게임 사용자 이탈을 예측하는데 미치는 영향을 밝혀내고자 한다. 더불어 실제 개발사의 패치 업데이트를 통해 본 연구에서 도출한 품질 불만 요인에 따라 어떻게 사용자수가 변화하는지 검증하고

자 한다. 이를 통해, 본 연구는 게임 개발사에게 게임 사용자 이탈 및 게임 품질 개선과 관련한 중요한 요인들을 도출하는 방법을 제시하고자 한다.

III. 연구 대상 선정 및 실증연구를 위한 데이터셋 구성

본 연구 대상으로 ㈜크래프톤의 “PLAYERUNKNOWN’S BATTLEGROUNDS”(이하 배틀그라운드, PUBG)를 선정하여 분석을 진행하였다. PUBG는 출시와 동시에 스팀 판매량 1위, 최고 사용자 수 300만 명을 달성(2017.11.27 기준)한 국내산 흥행 게임이다. 하지만, 출시 초기의 인기에도 불구하고, 최근 이용자 수가 크게 감소하는 추세를 보였다. 스팀 통계에 따르면 배틀그라운드의 동시 접속 사용자 수 (Peak Concurrent User)는 2018년 9월 100만 명 이하로 떨어진 이후에도 지속적인 하락세를 보이고 있다. 이는 배틀그라운드가 게임 서비스에 대한 사용자들의 의견을 적절히 피드백하지 못하여 이용자들이 하여금 게임을 이탈할 유인을 제공하였기 때문이다(유일환, 2018; 이재영, 2018). 배틀그라운드는 multi-player game(다인용 게임)으로, 하나의 게임 플레이 당 100명의 사용자가 같은 공간에서 동시에 플레이한다. 즉, 1회의 게임 플레이를 위해서는 100명의 사용자가 동시에 배틀그라운드를 플레이 해

야 한다는 것이다. 따라서, 배틀그라운드 게임의 생존에 있어 동시 접속 사용자 수 혹은 사용자 잔존율은 매우 중요한 요소라 할 수 있다.

본 연구의 실증 분석은 사용자들의 게임 리뷰 내 불만요소와 사용자 이탈과의 관계 확인을 위한 생존분석과, 사용자 수 추세변화 확인을 위한 시계열분석으로 구성된다. 이를 위해 각 분석의 특성에 맞춰 2개의 데이터셋을 구성하였다. 먼저 생존 분석을 위해 게임 관련 ESD(Electronic Software Distribution) 플랫폼인 스팀(Steampowered.com)에서 게임에 대한 사용자들의 리뷰 텍스트 데이터와 사용자들의 플랫폼 이용 행태 데이터를 수집하였다. 또한 준실험설계 분석을 위한 배틀그라운드 게임 사용자 수 관련 시계열 데이터는 게임 통계를 전문적으로 제공하는 스팀스파이(Steamspy.com) 플랫폼에서 수집하였다. 데이터의 수집 기간은 게임통계 데이터의 경우 2017년 11월 27일부터 2018년 11월 26일까지이며, 스팀 리뷰 데이터의 경우 2018년 2월 14일부터 2018년 3월 31일까지로 설정하여 수집하였다.

스팀 ESD 사용행태 데이터의 변수명과 설명은 <표 1>에서, 기술 통계는 <표 2>에서 확인할 수 있다. 또한, 게임 통계 시계열 데이터의 변수와 설명은 <표 3>, 기술 통계는 <표 4>, 상관관계 분석 결과는 <표 5>와 같다.

<표 1> ESD 사용행태 데이터 변수 설명

변수명	변수설명
NumTitlesOwned	사용자가 스팀 플랫폼 상에서 보유한 게임 개수(O’Neill <i>et al.</i> , 2016)
NumReviews	사용자가 스팀 플랫폼 상에서 작성한 리뷰 개수(Lin <i>et al.</i> , 2018)
TotalPlaytime	사용자의 총 누적 PUBG 플레이 시간(분)
2WeeksPlaytime	사용자의 게임 이탈 전 최근 2주간 PUBG 플레이 시간(분)(Alexandrovsky <i>et al.</i> , 2019)
PaidForPUBG	PUBG 게임 타이틀을 유상으로 구매했는지 여부
SteamPurchase	PUBG 외에 스팀에서 유상으로 구매한 다른 게임 타이틀 존재 여부
VotesUp	사용자가 PUBG에 대해 남긴 리뷰가 도움이 된다고 평가받은 횟수(Kang <i>et al.</i> , 2013)
VotesScore	사용자가 PUBG에 대해 남긴 리뷰의 점수(스팀 내부 알고리즘 추산)(Sifa <i>et al.</i> , 2015)

〈표 2〉 ESD 사용행태 데이터 기술 통계

변수명	Obs.	Mean	Std. Dev.	Min	Max
NumTitlesOwned	35,991	72	147	0	9,362
NumReviews	35,991	4	10	1	724
TotalPlaytime	35,991	12,154	13,661	5	138,732
2WeeksPlaytime	35,991	1,267	1,389	0	19,562
PaidForPUBG	35,991	0.956	0.205	0	1
SteamPurchase	35,991	0.766	0.423	0	1
VotesUp	35,991	0.514	1.070	0	41
VotesScore	35,991	0.182	0.243	0.000	0.756

〈표 3〉 게임 통계 시계열 데이터 변수 설명

변수명	변수 설명
PCCU	PUBG를 플레이한 일별 최고 동시접속자 수(Prodan and Nae, 2009)
PosReviews	스팀 플랫폼 상 PUBG 관련 일별 긍정적인 리뷰 수(Chen and Xie, 2008; Rezaei and Ghodsi, 2014)
NegReviews	스팀 플랫폼 상 PUBG 관련 일별 부정적인 리뷰 수(Chen and Xie, 2008; Rezaei and Ghodsi, 2014)
Event	PUBG 게임 내 이벤트 진행 여부(예: SNS 친구 추천 등)
StreamViewers	PUBG 관련 유튜브, 트위치 스트리밍 조회수(Sjöblom and Hamari, 2017)
NumOwners	스팀 플랫폼 상 PUBG 게임 타이틀 소유자 수

〈표 4〉 게임 통계 시계열 데이터 기술 통계

변수명	Obs.	Mean	Std. Dev.	Min	Max
PCCU	520	231,639	133,042	18,884	503,323
PosReviews	520	189.788	233.257	0.000	3,862.093
NegReviews	520	140.530	161.218	0.000	1,564.321
Event	520	0.260	0.439	0.000	1.000
StreamViewers	520	27,361	16,126	6,000	120,324
NumOwners	520	24,972,298	18,211,547	403,000	60,980,000

〈표 5〉 게임 통계 시계열 데이터 상관관계 분석

	(1)	(2)	(3)	(4)	(5)	(6)
(1) PCCU	1.000					
(2) PosReviews	0.056	1.000				
(3) NegReviews	0.374	0.330	1.000			
(4) Event	-0.020	-0.100	-0.042	1.000		
(5) StreamViewers	-0.448	0.165	-0.107	-0.149	1.000	
(6) NumOwners	0.236	-0.282	-0.046	0.297	-0.719	1.000

IV. 실증 분석

본 연구는 게임 사용자 불만사항에 대응하여 개발사가 패치를 통해 해당 게임을 개선하였을 때, 사용자 이탈을 방지할 수 있는지 확인하는 것을 주요 연구 질문으로 삼는다. 본 연구는 해당 연구 질문을 세 단계로 나누어 접근한다. 첫째, 사용자가 플랫폼에 남긴 게임 리뷰 데이터에 대한 텍스트 마이닝 분석을 수행하여 불만사항(혹은 요구사항) 요소를 식별한다. 둘째, 생존분석(Survival Analysis)을 통해 각 불만사항이 사용자의 게임 이탈 확률에 미치는 영향을 확인한다. 셋째, 준실험설계 방법론을 활용하여 사용자 불만사항에 대한 개발사의 대응(소프트웨어 개선 패치)이 실제 게임 사용자 수에 어떻게 영향을 미치는지 검증한다.

4.1 텍스트 마이닝 방법론을 활용한 사용자 리뷰의 주요 불만요소 추출

먼저 ‘스팀(Steam)’ 플랫폼에서 수집한 게임 사용자들의 게임 리뷰 데이터를 분석하여, 불만사항

(혹은 개선사항) 요소를 식별한다. 이를 위해 사용자 리뷰 데이터에 텍스트 마이닝 분석을 수행한다. 사용자 리뷰는 텍스트로 구성된 비구조화 데이터 형태를 취한다. 따라서, 리뷰 데이터를 분석 가능한 형태(구조화 데이터)로 변환하기 위해 TF-IDF (Term Frequency-Inverse Document Frequency) 방법론을 활용한다(Jones, 1972; Luhn, 1957). TF-IDF는 문서 내 등장 단어에 대해 가중치를 할당하는 분석 기법이다. 한 문서(단어의 집합) 내에서 특정한 단어의 빈도가 높을수록(Term Frequency), 전체 문서 집합에서 해당 단어가 적게 등장할수록(Inverse Document Frequency) 해당 단어에 큰 가중치를 할당한다. 본 연구에서는 전체 사용자 리뷰에 대해 TF-IDF를 계산하여, 가중치 상위 100개의 주요 불만사항(혹은 개선 요구사항) 키워드를 식별하였다. 본 연구에서는 요사이 자주 사용되고 있는 토픽 모델링보다는 TF-IDF 방법론을 활용하여 군집을 분석하였다. 토픽 모델링의 경우, 자동으로 토픽을 정해주고 해당 토픽에 대한 키워드를 파악할 수 있다. 그러나, 본 연구의 목적은 사용자 리뷰의 불만 요소를 미리 정하지 않고 자유롭게 도출하고자 하였다.

〈표 6〉 불만요소 설명과 관련 키워드

불만요소	대분류	설명	관련 키워드
PCF1: Social	Enjoyment	채팅, 커뮤니티 등 게임 내 사회 활동을 위한 시스템 빈약	community, chat, talk, friends, ...
PCF2: Balance		게임의 규칙 및 설계가 불균형해 편향된 플레이 유도	damage, chance, balance, find, ...
PCF3: Contents		다양한 맵, 아이템 등의 부재로 게임 플레이의 선택지 부족	vehicle, weapon, map, mode ...
PCF4: Suspense		게임의 규칙 및 설계가 지루한 게임플레이를 유도	huge, tedious, far, loose, ...
PCF5: Abuse		불법프로그램 예방, 핵 사용자의 제재 등, 게임 보안 빈약	report, hack, cheater, ban, ...
PCF6: Waiting	Convenience	비효율적인 매칭 시스템으로 인해 게임 플레이 전 대기시간 발생	waiting, long, alone, time, ...
PCF7: Bug		게임 소프트웨어에 버그 존재 (예: 비상식적 물리법칙)	bug, improve, error, care, ...
PCF8: UI/UX		UI가 직관적이지 않고 설정 및 게임 종류 선택 제약	direction, button, screen, ...
PCF9: Customizing		게임 내 스킨 등 미적 커스터마이징 요소 부족	skin, cash, box, customize, ...
PCF10: Optimization		서버, 소프트웨어 등 최적화가 이루어 지지 않아 지연현상 발생	ping, lag, base, server, ram, ...

따라서 토픽 모델링 보다는 TF-IDF 가 보다 적절한 방법론이라 판단하였다.

그 다음 사용자 리뷰의 주요 키워드로부터 사용자들의 다양한 불만사항을 식별하고 유사한 의미를 갖는 불만사항 키워드를 군집화하였다. 이하 본 연구에서는 각 불만사항 키워드들의 군집을 “불만요소(PCF, Player Complaint Factor)”로 명칭한다. 결과적으로 스팀 플랫폼 상에서 배틀그라운드 게임에 대해 높은 빈도로 논의되는 10개 불만사항 요소를 도출하였으며, 각 불만요소에 대한 설명과 연관 불만사항 키워드는 <표 6>과 같다. 분석 편의를 위해 크게 게임의 재미(Enjoyment)와 관련된 요소들과 게임 사용의 편의성(Convenience) 으로 대분류 하였다.

4.2 Cox 모델을 활용한 사용자 불만요소와 게임 이탈율의 관계 분석

본 연구에서는 사용자의 리뷰에서 식별된 각 불만요소가 사용자 이탈에 미치는 영향을 분석하고자 한다. 사용자 이탈 확률을 추정할 수 있는 다양한 모델(probit, logit, poisson, linear probability 등) 가운데, 본 연구는 단순한 이탈여부 뿐만 아니라 이탈까지 소요되는 시간적 요소를 함께 고려할 수 있는 Cox 회귀모델을 도입한다(Cox, 1972; Wong, 2011). Cox 모델은 특정 사건이 특정 시기에 발생할 위험도(hazard ratio)에 대해 설명변수들이 미치는 영향을 확률적으로 분석한다. 또한 Cox 모델은 추정에 사용되는 변수들의 분포에(log normal, Weibull 등) 제한을 두지 않는 장점을 지닌다. 본 연구의 Cox 회귀 모형은 다음과 같다:

$$h_i(t|X_i, Z) = \exp(\beta X_i + \delta Z) \times h_0(t)$$

여기서 X_i 와 Z 는 각각 다음과 같다:

$$X_i = [PCF1_i, PCF2_i, PCF3_i, \dots, PCF10_i] \\ \&$$

$$Z = [TotalPlaytime, 2WeeksPlaytime, \\ NumTitlesOwned, SteamPurchase, \\ PaidForPUBG, NumReviews, VotesUp, \\ VotesScore]$$

각 사용자에게 대하여, 연구 데이터의 마지막 관측 시점으로부터 14일 이내 게임 플레이 기록이 없다면 이를 사용자 이탈(churn)로 정의한다.¹⁾ 각 사용자의 이탈여부와 더불어 첫 플레이 날짜와 마지막 플레이 날짜 사이의 기간을 기록하여 모델의 종속변수(누적 위험도 cumulative hazard rates) 계산에 사용한다.

$h_0(t)$ 는 기저 위험도(baseline hazard)를 나타내고, X_i 는 사용자 i 에 대하여, 각각 리뷰를 통해 제시된 불만 요소를 나타내며, Z_i 는 ESD 플랫폼 이용 행태 측면의 통제 변수들로 이루어진 벡터이다. β 와 δ 는 이 변수들에 대한 최대 우도 추정치(MPLE)를 가지며, 각 변수의 이탈 위험에 미치는 영향의 크기와 방향을 나타낸다. 설명변수 X_i 는 사용자 리뷰 텍스트 마이닝을 통해 식별한 10개의 사용자 불만 요소로 구성된다. Cox 모델 사용 시, 사용자 이탈 위험도는 시간 t 가 증가함에 따라 지수적으로 증가한다. 따라서, 모델 추정의 결과는, 사용자가 리뷰로 불만사항을 제시한 이후 t 시점에서 불만사항이 없는 경우보다 $\exp(\beta X_i)^t$ 배의 이탈확률을 가진다고 해석할 수 있다.

생존분석 모델의 추정 결과는 <표 7>에서 확인할 수 있다. 모델의 일관성을 확인하기 위해 주요 불만요소 변수를 제외한 모델 1과 불만요소를 포함한 모델 2를 나누어 추정하였다. 통제변수가 포함된 연구 모형(모델 2) 역시 일관되게 결과가 추정되었다. 따라서, 이하 모델 추론 결과에 대한 논의는 모두 모델 2의 결과를 사용하였다.

¹⁾ 본 연구 데이터를 통해 확인한 결과, 사용자가 게임을 플레이 후 10일 이상 경과하면 사용자 재방문 확률이 급격히 감소하고 다시는 게임으로 돌아오지 않았다. 이에 본 연구에서는 이탈의 정의에 따른 분석 결과의 민감도를 낮추기 위해 10일보다 보수적인 14일의 비활동기간을 사용자 이탈의 조건으로 설정하였다.

〈표 7〉 Cox 모델 추론 결과

	Model 1		Model 2	
	Coef	Exp(Coef)	Coef	Exp(Coef)
Player Complaint				
PCF1: Social			0.0218 (0.0308)	1.0221
PCF2: Balance			0.1000 ^{***} (0.0175)	1.1052
PCF3: Contents			0.1194 ^{***} (0.022)	1.1268
PCF4: Suspense			0.2024 ^{***} (0.0266)	1.2243
PCF5: Abuse			0.1020 ^{***} (0.0227)	1.1074
PCF6: Waiting			-0.0084 (0.0301)	0.9916
PCF7: Bug			0.0732 ^{**} (0.0298)	1.0759
PCF8: UI/UX			0.0516 [*] (0.0265)	1.0529
PCF9: Customizing			0.0715 ^{***} (0.0156)	1.0741
PCF10: Optimization			0.0700 ^{***} (0.0223)	1.0726
Playtime				
TotalPlaytime	0.0001 ^{***} (0.0000)	1.0000	0.0001 ^{**} (0.0000)	1.0000
2WeeksPlaytime	-0.0003 ^{***} (0.0000)	0.9997	-0.0002 ^{***} (0.0000)	0.9998
ESD Platform Use				
NumTitlesOwned	-0.0010 ^{***} (0.0002)	0.9990	-0.0003 (0.0002)	0.9997
SteamPurchase	-0.1349 ^{***} (0.0153)	0.8738	-0.1286 ^{***} (0.0212)	0.8794
PaidForPUBG	-0.1227 ^{***} (0.0378)	0.8845	-0.3898 ^{***} (0.0427)	0.6772
NumReviews	0.0118 ^{***} (0.0007)	1.0119	0.0101 ^{***} (0.0014)	1.0102
VotesUp	-0.0571 ^{***} (0.0073)	0.9445	-0.0746 ^{***} (0.0229)	0.9281
VotesScore	0.1742 (0.2990)	1.1903	-0.2196 ^{***} (0.0532)	0.8029
Likelihood ratio test	175.0 ^{***}		56.2 ^{***}	
#Observations	31,850		31,850	
# Events (player churn)	26,754		26,754	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard errors are in parentheses.

주요 불만요소 변수의 경우 대부분 통계적으로 유의미한 것으로 확인되었다. 다만 PCF1과 PCF6의 경우 모두 통계적으로 유의미하지 않다. 하지만, 그 외 불만요소 변수는 모두 양수계수를 가지며 통계적으로 유의미하다.

세부적으로는 불만요소 중 PCF4: Suspense이 가장 큰 계수(0.2024)를 가지며 PCF3: Contents, PCF5: Abuse, PCF2: Balance 순서로 큰 계수(0.1194, 0.1020, 0.1000)를 보인다. 상대적으로 이탈위험이 큰 불만요소들(PCF2, 3, 4, 5)은 게임의 재미(Enjoyment)와 연관성이 깊다(Klimmt *et al.*, 2009; Trepte and Reinecke, 2010). 지수변환을 통해 확인하면, PCF4의 경우 하루에 약 22.4% 씩 사용자 이탈 위험을 높인다. PCF3, PCF5, PCF2의 경우 하루 평균 이탈 위험 증가율이 12.7%, 10.7%, 10.5%로 확인된다.

또한 PCF7: Bug, PCF8: UI/UX, PCF9: Customizing, PCF10: Optimization의 경우 상대적으로 위험도가 낮게(추정치: 0.0732, 0.0516, 0.0715, 0.0700; 유의성: $p < 0.01$, $p < 0.05$, $p < 0.001$, $p < 0.001$) 추정되었다. 추정치를 지수 변환하면 이들 불만요소들은 각 7.6%, 5.3%, 7.4%, 7.3%의 일평균 이탈 위험도 증가에 기여한다. 이탈위험이 작은 불만요소들(PCF7, 8, 9, 10)은 게임 사용의 편의성과 관련되어 있는 것을 확인하였다(Chen and Zarki, 2011). 즉, 재미와 사용 편의성 두 측면이 모두 게임 개발에 있어 중요한 요소로 작용하지만, 전자가 후자보다 사용자 이탈에 더 큰 영향을 미치는 것으로 나타났다.

본 연구에서는 ESD 플랫폼 이용 행태와 게임 플레이 행태 등을 통제변수로 사용하였다. 따라서, Cox 모델 추정 결과 중 통제 변수들이 사용자 이탈 위험도에 미치는 영향 역시 살펴보았다. 게임 플레이 행태를 나타내는 변수로는 TotalPlaytime(총 플레이 시간)과 2WeeksPlaytime(최근 2주간 플레이 시간)이 있으며 각각 0.0001($p < 0.01$), -0.0002($p < 0.01$)의 이탈 위험도 추정 결과를 가진다. 총 플레이 시간(TotalPlaytime)은 양의 계수를

가져 사용자의 이탈 위험을 증가시킨다. 반면 최근 2주간 플레이 시간(2WeeksPlaytime)이 많을수록 이탈확률을 감소시키는 것을 알 수 있다. 이는 기존 게임 이탈 문헌에서 간과된 측면으로, 게임 플레이 시간을 측정함에 있어 전체 누적 기간과 최근 기간을 나누어 분석에 활용해야 한다는 시사점을 가진다(Alexandrovsky *et al.*, 2019). 끝으로 본 연구는 ESD 플랫폼 사용 행태와 사용자 이탈을 연관 지어 새로운 측면의 인사이트를 제공할 수 있다. 대표적으로 플랫폼 상에서 실제 지출이 발생한 경우(SteamPurchase, PaidForPUBG가 참인 경우) 게임 이탈 위험도를 일평균 40.4% 가량 감소시키는 것으로 나타났다. 또한, 리뷰를 많이 작성하는 것(NumReviews)은 사용자 이탈 위험을 증가시키지만, 리뷰가 플랫폼에 의해 높게 평가(VotesScore)되거나 많은 관심(VotesUp)을 받으면 게임 이탈 위험을 크게 감소시킨다(추정치: -0.2196, -0.0746; 유의성: $p < 0.01$, $p < 0.01$). 이는 사용자의 리뷰가 많은 관심을 받으면, 다음 게임 패치 때 해당 불만사항이 반영될 확률이 높아지기 때문으로 추측된다(Füller *et al.*, 2008).

4.3 준실험설계 방법론을 활용한 개발사의 불만요소 대응 시 사용자 증감을 비교

앞서 본 연구에서는 텍스트마이닝과 생존분석 방법론을 활용하여 사용자의 리뷰에 기재된 주요 불만사항을 식별하고, 관련 불만 요소가 사용자 이탈에 미치는 영향을 분석하였다. 본 절에서는 준실험설계 방법론을 도입하여 상기 분석의 결과를 검증하고자 한다.

즉, 본 연구 결과에 따라 게임 개발사는 사용자의 불만사항들을 해결(또는 개선)하는 게임 패치를 통해 사용자 이탈을 방지할 수 있을 것이다. 이를 검증하기 위한 준실험설계 연구 방법은 다음과 같다. 1) 배틀그라운드 개발사가 배포한 패치 노트를 통해, 사용자 불만요소가 반영된 게임 패치를 확인한다. 또한 사용자 이탈의 정의에 따라, 이

전 혹은 이후 패치와의 간격이 2주 미만인 패치는 데이터에서 제외한다. 2) Cox 모델 추정 결과 검증 을 위해, 사용자 이탈 위험이 높은 불만요소(PCF4: Suspense, PCF3: Contents, PCF5Abuse, PCF2: Balance)가 반영된 패치와 위험도가 낮은 불만요소(PCF7: Bug, PCF8: UI/UX, PCF9: Customizing, PCF10: Optimization)를 해결한 패치를 구분하였다. 이 때, 각 패치가 개선한 불만요소별 Cox 모델의 이탈위험도 추정값을 다음과 같이 연산하여 각 패치를 통해 기대되는 총 이탈 위험 감소율을 계산한다: $[\exp(\beta X_i)]^{-1}$.²⁾ 결과적으로, 첫 번째 패치의 위험도 감소율은 46.02%, 두 번째 패치의 위험도 감소율은 25.64%로 추정된다(PUBG, 2018a, 2018b). 3) 불만요소를 개선하는 패치가 이루어지기 전과 패치가 이루어진 후의 사용자 수 시계열 특성을 분석하여 전후의 차이가 존재하는지 여부를 확인한다. 앞서 계산한 이탈 위험 감소율과 사용자 수 시계열 분석 결과가 비슷하게 제시된다면, Cox 모델로 확인한 주요 불만요소의 개선이 실제 사용자 숫자 증가로 이어질 수 있다는 근거로 활용될 수 있을 것이다.

사용자의 이탈 위험도는 시간이 흐름에 따라 지수적으로 증가하는 값이다. 따라서, 시계열 데이터에서 게임 패치 전과 후의 사용자 이탈율의 차이를 보려면 사용자 수의 평균값이 아닌 증감 추세에 대한 분석이 필요하다. 본 연구는 개입(Intervention)모형 중 시계열 간 추세 차이 비교가 용이한 간역시계열(Interrupted Time Series; ITS)설계를 활용한다. 준실험설계 모형은 다음과 같다.

$$\begin{aligned}
 PCCU_t = & \alpha + \beta_1 Date_t + \beta_2 AfterPatch_t \\
 & + \beta_3 Date_t \times AfterPatch_t \\
 & + \delta_1 PosReviews_t + \delta_2 \sim Reviews_t \\
 & + \delta_3 Event_t + \delta_4 StreamViewers_t \\
 & + \delta_5 NmOwners_t + \epsilon_t
 \end{aligned}$$

²⁾ 패치에서 개선한 불만요소들의 추정된 위험도를 지수변환한 후 그 역수를 누적 곱함.

모형에는 시간변수 $Date_t$ 를 포함한다. 추정값 β_1 은 시계열 데이터의 추세(trend)를 나타낸다. 또한, 본 모형은 더미변수 $AfterPatch_t$ 를 통해 게임 패치 전(“”)과 후(“1”)의 시점을 구분 짓는다. 따라서, 두 변수의 상호작용항 $Date_t \times AfterPatch_t$ 의 추정치 β_3 는 두 시점 사이의 추세 차이를 나타내며, 이는 본 준실험설계를 통해 밝히고자 하는 게임 패치의 효과를 대표하는 값이다.

본 연구에서는 GLS(Generalized Least Squares)를 활용해 상기 모형을 추정한다. GLS모형의 자기상관(AR) 구조는 ARIMA 파라미터를 탐색하여 가정한다(Hyndman and Khandakar, 2008). 본 데이터는 ARIMA(5, 1, 1) 추정에서 AR(5)의 계수들이 0.2에 수렴하여 GLS 추정에 continuous autoregressive process를 가정하고 0.2의 AR 상수를 사용하였다(Hansen, 2007). 모델 추정에는 각 패치 일자를 기준으로 시계열 데이터를 패치 전후 14일씩(총 28일)의 데이터만 남겨 총 2개의 데이터셋을 구성하여 사용하였다.

<표 8>에서 ITS 설계의 GLS 추정 결과를 확인할 수 있다. Date의 계수가 음수라는 점을 고려하였을 때, 게임은 이미 사용자 감소 추세를 경험하고 있는 상황이다. ITS로 확인하고자 한 패치 전후 추세 차이($Date \times AfterPatch$ 항의 추정치)는 두 패치 모두 양수 계수(추정치: 4,555, 1,833; 유의성: $p < 0.05$, $p < 0.01$)를 가진다. 이를 통해 불만요소를 개선하는 패치가 실제로 사용자 수 감소 추세를 완화하고 게임 사용자 수에 긍정적인 영향을 미치는 것을 알 수 있다.

또한, 각 계수의 비교를 위해 모든 변수를 표준화하여 추정한 값도 <표 8>에 제시하였다. 각 패치의 z-value(Standardized Coefficient) 컬럼을 살펴보면 패치A의 ITS 추세 차이가 0.953으로 패치B의 추세 차이, 0.393, 보다 2배가량 더 크다. 이 같은 패턴은 각 패치의 이탈 위험 감소율에서도 동일하게 나타난다. 패치 A의 이탈 위험도 감소율(46%) 역시 패치 B의 위험도 감소율(26%)보다 약 2배 더 큰 것을 앞서 추정하였다. 따라서, 이탈 위

〈표 8〉 ITS 설계의 GLS 추정 결과

Hazard Reduction	Patch A		Patch B	
	46%		26%	
DV: PCCU	Estimates	z-value	Estimates	z-value
Date	-137 (442)	0.572	-738 (468)	0.070
AfterPatch	-80,636,898** (30,423,783)	-0.738	-32,432,249 (10,915,526)	-0.184
Date x AfterPatch	4,555** (1,718)	0.953	1,833*** (617)	0.393
PosReviews	66.00* (34.00)	0.149	49.00* (27.00)	0.076
NegReviews	-28.00 (40.00)	-0.101	-8.00 (28.00)	-0.018
Event	-5,921 (4,745)	-0.102	-8,112** (3,653)	-0.098
StreamViewers	1.000** (0.000)	0.182	1.000*** (0.000)	0.159
NumOwners	0.000*** (0.000)	-1.091	0.000*** (0.000)	-0.819
No.of Dates	28		28	
AIC	912.995		1,267.474	
Log Likelihood	-445.498		-622.737	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard errors are in parentheses.

험 감소율과 ITS 추세 차이에서 일관된 패턴이 확인된다.

즉, ITS 준실험 설계 결과를 통해, 앞선 텍스트 마이닝과 생존분석 결과를 실증적으로 뒷받침할 수 있었다. 결과적으로 개발사는 게임 패치를 기획함에 있어 사용자의 의견(리뷰)을 수렴하고, 그 중 중요도가 더 높은 불만요소를 우선적으로 개선한다면 사용자 이탈의 추세를 획기적으로 낮출 수 있을 것이다.

V. 결 론

본 연구는 사용자 불만 요인이 게임 사용자 이탈에 미치는 영향에 대해 연구하였다. 또한 준실험설계 방법론을 통해 본 연구는 게임 불만 사항 개선이

사용자의 이탈에 유의미한 영향을 미친다는 것을 확인했다. 선행 연구에서는 주로 사용자의 게임 참여 동기 관점에서 이탈 요인을 살펴보았으나, 본 연구는 반대 개념인 사용자들의 불만요인에 따른 사용자 이탈 여부를 분석하였다. 특히, 게임 리뷰 데이터를 통해 직접적으로 사용자 불만 요인을 도출하였고, 이 불만 요인에 따라 사용자 이탈에 미치는 영향을 실증적으로 보였다. 이에 따라 향후 리뷰 데이터가 게임품질 불만족을 파악할 수 있는 지표로 활용될 수 있을 것이다. 또한 본 연구에서는 텍스트 마이닝, Cox 모델, 준실험설계 연구방법을 결합한 새로운 방법론을 제시함에 따라 의미가 있다고 할 수 있다. 이는 새로운 방법론으로의 확장이라는 측면과, 게임 산업 내에 새로운 프레임워크를 제시하였다는 점에서 의의를 갖는다.

5.1 연구결과의 학문적 기여

본 연구는 사용자 불만 요인이 사용자 이탈에 미치는 영향을 실증 연구하였다는 점에서 학문적 의의를 갖는다. 기존의 연구들은 설문을 통해 개별 사용자의 게임 참여 동기와 인구통계학적 특성을 통해 사용자의 이탈예측을 밝혀내고자 하였다. 하지만, 사용자가 느끼는 게임에 대한 불만 요인과 그 개선에 대한 특성을 간과하였다. 게임은 하나의 제품으로서 각각의 서비스 품질 특성을 갖고 있다. 따라서 불만 사항을 개선한 게임 품질은 사용자 이탈 방지의 요인으로 작용할 수 있다. 이러한 점에 착안해 본 연구는 게임 불만 요인이 사용자의 이탈에 미치는 영향을 살펴보았다. 이에 더 나아가 게임의 불만 요인에 대한 개선여부에 따른 사용자 추이를 살펴봄으로써, 게임의 불만 요인이 게임의 생존과 직결되어 있다는 것을 밝힌 첫 연구로써 그 의미를 갖는다.

또한 본 연구는 다양한 방법론의 결합을 통해 방법론의 확장을 시도하였다는 점에서 학문적으로 기여한다. 즉, 머신러닝 기법인 텍스트 마이닝과 계량경제학 모델인 Cox Regression Model을 결합하여 분석 방법론을 제시하였다. 1) TF-IDF 가중치를 활용하여 사용자 리뷰 데이터로부터 게임 불만 요인을 추출해, 이를 Cox Regression Model의 설명변수로 활용하였다. 2) 이후 간여시계열(Interrupted Time Series; ITS) 방법론을 활용하여 준실험설계를 수행하였다. 3) 이를 통해 불만 요인 개선에 따른 사용자 이탈 추이를 제시함으로써 본 연구 방법론을 검증하였다. 즉, 본 연구는 머신러닝, 계량경제학 그리고 실험방법론을 모두 활용하여, 게임 사용자 이탈을 연구하는 문헌에 새로운 분석 방법론을 제시하였다는 점에서 의의를 가질 수 있다.

5.2 연구결과의 실증적 기여

본 연구의 실증적 기여는 다음과 같다. 게임이 고부가가치 제품이라는 특성이 나타나기 위해서

는 고객 불만 사항을 반영한 게임 품질의 개선이 필수적으로 고려되어야 한다는 것을 제시한다. 장기적인 수익창출을 위해 개발사는 사용자가 지속적으로 게임 서비스를 이용할 수 있도록 유인을 제공해야한다. 본 연구는 게임 불만 요인이 사용자 이탈 방지의 중요 요소임을 실증적으로 보이고, 이를 통해 게임 사용자 유지 관리가 가능하다는 것을 제시한다. 또한 본 연구는 게임산업 내에서 고객 리뷰가 의사결정 지원 도구로 활용될 수 있음을 밝혀, 게임 개발자 및 관리자의 전략적 의사결정 도구를 다양화한다는 점에서 기여한다. 리뷰는 고객의 의사를 전달하는 효과적인 VoC(voice-of-customer) 통로로 제품 및 서비스의 질, 고객의 불만사항 및 선호와 필요 등을 직접적으로 발견할 수 있는 수단이다. 하지만, 기존의 연구들에서는 리뷰를 이탈의 주요한 예측변수로 다루지 않았다. 따라서 본 연구는 다양한 방법론을 적용하여, 사용자 리뷰로부터 사용자들이 중요하게 인지하는 불만 요인을 도출하였고, 이들이 사용자 이탈에 미치는 영향을 밝혀냈다. 이를 통해 게임 산업 내에서 전반적인 의사결정에 온라인 리뷰를 전략적으로 활용할 수 있음을 제시하였다.

마지막으로 새로운 방법론을 제시함으로써 사용자 이탈 예측 및 방지를 위한 새로운 분석 프레임워크를 제시한다. 사용자 리뷰 수집과 Text Mining, Cox Model, 그리고 Quasi-experiment 분석 기법 등을 통해 개발사는 중요 불만 요인을 도출하고, 의사결정에 있어 우선순위를 설정할 수 있다. 이는 게임 개발사뿐만 아니라 콘텐츠 개발 업체 및 콘텐츠 서비스 제공업체들의 고객 이탈 방지 예측 및 서비스 업데이트를 위한 의사결정에도 도움을 줄 수 있다. 더불어 이는 제품 수명 연장을 위해 활용될 수 있을 것이다.

5.3 연구의 한계

본 연구는 사용자 불만 요인이 사용자이탈에 미치는 영향을 실증적으로 연구한 최초의 연구라

는 점에서 의의를 갖는다. 그럼에도 불구하고 일부 한계점을 갖는다. 먼저 자료수집과 이용에 있어, 영어와 한국어로 작성된 리뷰만을 수집하여 이용하였다. 이는 Text Mining 분석기법에 의한 한계로 다른 언어의 경우 분석에 어려움이 있어 본 연구에서는 반영되지 않았다. 다음으로 해당 연구는 멀티플레이를 지원하는 온라인 게임을 대상으로 진행되었기 때문에 연구결과가 멀티플레이를 지원하지 않는 콘솔게임에도 일반적으로 적용되기에는 무리가 있다. 이와 함께 배틀그라운드라는 특정 게임에 대해 실증 분석을 진행하였기에, 다른 종류의 멀티플레이 지원 게임을 대상으로 검증할 필요가 있다. 배틀그라운드가 현재 전 세계적으로 인기있는 게임으로 어느 정도 멀티플레이 게임을 대표할 수 있을 것으로 판단되나, 후속 연구들은 보다 다양한 게임을 대상으로 본 연구의 방법론을 적용해 분석해 볼 필요가 있다.

참 고 문 헌

- [1] 게임메카, “[아뿔싸] 던파 사용자들, 강화 소실 방지 아이템에 ‘들썩’” 중앙일보, 2011.08.29., Available at <https://news.joins.com/article/6078972>.
- [2] 삼성KPMG, “게임 산업을 둘러싼 10대 변화 트렌드”, 2018, Issue Monitor, 89.
- [3] 유일환, “배틀그라운드, 스팀 동시 접속자 90만 밑으로 하락” 디스이즈게임, 2018.10.11., Available at http://www.thisisgame.com/webzine/news/nboard/4/?n=87394&utm_source=naver&utm_medium=outlink&utm_campaign=tigadmi n74&utm_content=87394&nogate.
- [4] 이재영, “분기 수익 ‘반 토막’ 펍지, 배틀그라운드에 무슨 일이?”, 인베스트조선, 2018.10.15., Available at <http://www.investchosun.com/2018/10/15/3231056>.
- [5] 한국콘텐츠진흥원, “2017년 4분기 콘텐츠산업 규모: 2017년 4분기 및 연간 콘텐츠산업 동향분석보고서”, 2018.
- [6] 한국콘텐츠진흥원, “2017~2018년 상반기 국내 게임 산업 10대 이슈. 대한민국 게임백서 2018”, 2019.
- [7] Alexandrovsky, D., M. A. Friehs, M. V. Birk, R. K. Yates, and R. L. Mandryk, “Game dynamics that support snacking, not feasting”, In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, October 2019, pp. 573-588.
- [8] Borbora, Z. H. and J. Srivastava, “User behavior modelling approach for churn prediction in online games”, *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*, 2012, pp. 51-60.
- [9] Castro, E. G. and M. Tsuzuki, “Churn prediction in online games using players’ login records: A frequency analysis approach”, *IEEE Transactions on Computational Intelligence and AI in Games*, Vol.7, No.3, 2015, pp. 1-1.
- [10] Chen, K. T., P. Huang, and C. L. Lei, “How sensitive are online gamers to network quality?”, *Communications of the ACM*, Vol.49, No.11, 2006a, pp. 34-38.
- [11] Chen, V. H. H., H. B. L. Duh, P. S. K. Phuah, and D. Z. Y. Lam, “Enjoyment or engagement? Role of social interaction in playing massively mulitplayer online role-playing games (MMORPGS)”, *International Conference on Entertainment Computing. ICEC 2006*, September 2006b, pp. 262-267.
- [12] Chen, P. and M. E. Zarki, “Perceptual view inconsistency: An objective evaluation framework for online game quality of experience”, *10th Annual Workshop on Network and Systems Support for Games*, 2011, pp. 1-6.
- [13] Chen, Y. and J. Xie, “Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix”, *Management Science*,

- Vol.54, No.3, 2008, pp. 477-491.
- [14] Cheung, C. M. K., X. L. Shen, Z. W. Y. Lee, and T. K. H. Chan, "Promoting sales of online games through customer engagement", *Electronic Commerce Research and Application*, Vol.14, No.4, 2015, pp. 241-250.
- [15] Cox, D. R., "Regression models and life tables (with discussion)", *Journal of the Royal Statistical Society: Series B*, Vol.34, No.2, 1972, pp. 187-202.
- [16] Evans, E., "The economics of free freemium games, branding and the impatience economy", *The International Journal of Research into New Media Technologies*, Vol.22, No.6, 2016, pp. 1-18.
- [17] Füller, J., K. Matzler, and M. Hoppe, "Brand community members as a source of innovation", *Journal of Product Innovation Management*, Vol.25, No.6, 2008, pp. 608-619.
- [18] Griffiths, M. D., M. N. O. Davies, and D. Chappell, "Demographic factors and playing variables in online computer gaming", *Cyber Psychology & Behavior*, Vol.7, No.4, 2004, pp. 479-487.
- [19] Hadji, F., R. Sifa, A. Drachen, C. Thureau, K. Kersting, and C. Bauckhage, "Predicting player churn in the wild", *2014 IEEE Conference on Computational Intelligence and Games*, August 2014, pp. 1-8.
- [20] Hansen, C. B., "Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects", *Journal of Econometrics*, Vol.140, No.2, 2007, pp. 670-694.
- [21] Hyndman, R. J. and Y. Khandakar, "Automatic time series forecasting: The forecast package for R", *Journal of Statistical Software*, Vol.27, No.3, 2008, pp. 1-22.
- [22] Jones, S. K., "A statistical interpretation of term specificity and its application in retrieval", *Journal of Documentation*, Vol.28, No.1, 1972, pp. 11-21.
- [23] Jung, H. S., K. H. Kim, and C. H. Lee, "Influences of perceived product innovation upon usage behavior for MMORPG: Product capability, technology capability, and user centered design", *Journal of Business Research*, Vol.67, No.10, 2014, pp. 2171-2178.
- [24] Kang, A. R., J. Y. Woo, J. Y. Park, and H. K. Kim, "Online game bot detection based on party-play log analysis", *Computers & Mathematics with Applications*, Vol.65, No.9, 2013, pp. 1384-1395.
- [25] Kawale, J., A. Pal, and J. Srivastava, "Churn prediction in MMORPGs: A social influence based approach", *2009 International Conference on Computational Science and Engineering*, Vol.4, 2009, pp. 423-428.
- [26] Klimmt, C., D. Hefner, and P. Vorderer, "The video game experience as 'true' identification: A theory of enjoyable alterations of players' self-perception", *Communication Theory*, Vol.19, No.4, 2009, pp. 351-373.
- [27] Lin, D., C. P. Bezemer, and A. E. Hassan, "An empirical study of early access games on the Steam platform", *Empir. Softw. Eng.*, Vol.23, No.2, 2018, pp. 771-799.
- [28] Luhn, H. P., "A statistical approach to mechanized encoding and searching of literary information", *IBM Journal of Research and Development*, Vol.1, No.4, 1957, pp. 309-317.
- [29] Martončík, M. and J. Lokš, "Do World of Warcraft (MMORPG) players experience less loneliness and social anxiety in online world (virtual environment) than in real world (offline)?", *Computers in Human Behavior*, Vol.56, 2016, pp. 127-134.
- [30] Milošević, M., N. Živić, and I. Andjelković, "Early churn prediction with personalized targeting in mobile social games", *Expert Systems with*

- Applications*, Vol.83, 2017, pp. 326-332.
- [31] Moon, J., M. D. Hossain, G. L. Sanders, E. J. Garrity, and S. Jo, "Player commitment to massively multiplayer online role-playing games (MMORPGs): An integrated model", *International Journal of Electronic Commerce*, Vol.17, No.4, 2013, pp. 7-38.
- [32] O'Neill, M., E. Vaziripour, J. Wu, and D. Zappala., "Condensing steam: Distilling the diversity of gamer behavior", *IMC '16 Proc. of the Internet Measurement Conference (IMC)*, ACM, Santa Monica, USA, 2016, pp. 81-95.
- [33] Parasuraman, A., V. A. Zeithaml, and L. L. Berry, "Servqual: A multiple-item scale for measuring consumer perc", *Journal of Retailing*, Vol.64, No.1, 1988, pp. 12-40.
- [34] Prodan, R. and V. Nae, "Prediction-based real-time resource provisioning for massively multiplayer online games", *Future Generation Computer Systems*, Vol.25, No.7, 2009, pp. 785-793.
- [35] PUBG, "PC 1.0 Update #13 - Playerunknown's battlegrounds. playerunknown's battlegrounds", 2018a, <https://www.pubg.com/2018/06/19/pc-1-0-update-13/>.
- [36] PUBG, "PC 1.0 Update #14 - Playerunknown's battlegrounds. playerunknown's battlegrounds", 2018b, <https://www.pubg.com/2018/06/19/pc-1-0-update-14/>.
- [37] Rezaei, S. and S. S. Ghodsi, "Does value matters in playing online game? An empirical study among massively multiplayer online role-playing games(MMORPGs)", *Computers in Human Behavior*, Vol.35, 2014, pp. 252-266.
- [38] Sifa, R., F. Hadiji, J. Runge, A. Drachen, K. Kersting, and C. Bauckhage, "Predicting purchase decisions in mobile free-to-play games", *In Proceeding of AAAI AIIDE*, 2015.
- [39] Sjöblom, M. and J. Hamari, "Why do people watch others play video games? An empirical study on the motivations of twitch users", *Computers in Human Behavior*, Vol.75. 2017, pp. 1-12.
- [40] Trepte, S. and L. Reinecke "Avatar creation and video game enjoyment effects of life-satisfaction, game competitiveness, and identification with the avatar", *Media Psychology*, Vol.22, No.4, 2010, pp. 171-184.
- [41] Wong, K. K. K., "Using cox regression to model customer time to churn in the wireless telecommunications industry", *Journal of Targeting, Measurement and Analysis for Marketing*, Vol.19, No.1, 2011, pp. 37-43.
- [42] Wu, J. H., S. C. Wang, and H. H. Tsai, "Falling in love with online games: The uses and gratifications perspective", *Computers in Human Behavior*, Vol.26, No.6, 2010, pp. 1862-1871.
- [42] Yee, N., "The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments", *Presence: Teleoperators and Virtual Environments*, Vol.15, No.3, 2006, pp. 309-329.

Estimate Customer Churn Rate with the Review-Feedback Process: Empirical Study with Text Mining, Econometrics, and Quai-Experiment Methodologies

Choi Kim^{*} · Jaemin Kim^{**} · Gahyung Jeong^{***} · Jaehong Park^{****}

Abstract

Obviating user churn is a prominent strategy to capitalize on online games, eluding the initial investments required for the development of another. Extant literature has examined factors that may induce user churn, mainly from perspectives of motives to play and game as a virtual society. However, such works largely dismiss the service aspects of online games. Dissatisfaction of user needs constitutes a crucial aspect for user churn, especially with online services where users expect a continuous improvement in service quality via software updates. Hence, we examine the relationship between a game's quality management and its user base. With text mining and survival analysis, we identify complaint factors that act as key predictors of user churn. Additionally, we find that enjoyment-related factors are greater threats to user base than usability-related ones. Furthermore, subsequent quasi-experiment shows that improvements in the complaint factors (i.e., via game patches) curb churn and foster user retention. Our results shed light on the responsive role of developers in retaining the user base of online games. Moreover, we provide practical insights for game operators, i.e., to identify and prioritize more perilous complaint factors in planning successive game patches.

Keywords: *Review-Feedback Process, Customer Churn, Game UserSatisfaction, Cox-harzard model, Quai-experiment*

* Master Student, Management Information Systems, Kyung Hee University

** Master student, Big Data Management, Kyung Hee University

*** Master student, Big Data Management, Kyung Hee University

**** Corresponding Author, Professor, Management, Kyung Hee University.

◎ 저 자 소 개 ◎



김 초 이 (lunalice73@khu.ac.kr)

경희대학교 경영학과에서 학부학위를 취득하였으며, 현재 동대학원 경영학과 석사 과정을 수료하였다. 관심분야는 온라인 게임, 웹툰 소비자 행태 연구이며, 관련하여 WeB 2019 학회에서 논문을 발표하였다.



김 재 민 (kimjaem@khu.ac.kr)

경희대학교 호텔관광대학에서 학부학위를 취득하였으며, 현재 동대학원 빅데이터 경영학과에서 석사과정 중이다.



정 가 현 (sky2193@khu.ac.kr)

경희대학교 경영학과에서 학부학위를 취득하였으며, 현재 동대학원 빅데이터 경영학과에서 석사과정 중이다.



박 재 흥 (jaehp@khu.ac.kr)

University of Texas at Austin에서 경영학 박사학위를 취득하였으며, 울산과기원을 거쳐 경희대에서 교수로 재직 중이다. Information System Researches 등에 계량모델링 및 머신러닝 방법을 활용한 다수의 논문을 게재 하였다. 현재 경희대학교 학생처장, 창업지원단장, 캠퍼스타운 단장을 역임 중에 있다.

논문접수일 : 2021년 05월 31일

게재확정일 : 2021년 06월 24일