



사용자 분석 입문 (Intro to User Analytics)

작성자: (Anders Drachen, Alessandro Canossa, Magy Seif El-Nasr)

작성일: 2013년 5월 30일

이 글은 가마수트라의 자매지인 게임 디벨로퍼 매거진(Game Developer magazine) 2013년 5월판에 실린 글로, 안드로이드 개발을 둘러싼 도전들과 기회들을 알아내기 위해 2013년 개발자들을 대상으로 실시한 설문조사를 담고 있다. [2013년 5월호 구입하기](#)¹

게임 분석은 최근 몇 년 사이 큰 주목을 받았다. 게임 개발 사이클에 분석학을 도입하는 것은 플레이어에 대한 더 좋은 지식을 얻으려는 필요에 의한 것이었다. 경영과 기획 등을 포함한 게임 회사의 많은 부서들에게 도움이 되기 때문이다. 따라서 게임 분석은 게임 산업에서 점점 더 중요한 비즈니스 인텔리전스(business intelligence)의 영역이 되고 있다. 원격 측정, 시장 보고, 품질 보증 시스템, 벤치마크 테스트, 그리고 그 밖의 많은 원천들을 통해 얻은 모든 양적인 데이터는 비즈니스 인텔리전스와 정보를 주는 의사 결정에 영향을 끼친다.

게임 개발 과정에 분석학을 도입할 때 가장 중요한 문제 두 가지는 무엇을 추적할 것인가와 어떻게 데이터를 분석할 것인가이다. 어떤 데이터를 모을 것인지를 결정하는 과정은 특징 선택(feature selection)이라고 한다. 특징 선택은 특히 그것이 사용자의 행동으로 이어질 때 도전이 된다. 우리가 어떤 행동을 추적할 것인지를 결정하기 위해 적용할 수 있는 단 하나의 올바른 대답이나 표준적인 모델은 없다. 그 대신 예를 들어 사용자의 경험을 향상시키거나 수익화(monetization)를 증가시키는 등 다양한 목표에 따라 다양한 전략이 가능하다. 이 글에서 우리는

¹ 참조링크: <http://gdmag.com/issue/2013/May>

특징 선택을 글 전체의 주제로 다루면서, 사용자 중심의 게임 분석의 몇 가지 근본적인 관심사들에 대해 대략적으로 다루고자 한다. 먼저 추적할 수 있는 사용자 데이터의 유형들에 대해 다루고 난 후, 개발자들이 무엇을 그리고 어떻게 측정할 것인지를 선택하는 특징 선택의 과정에 대해 다룰 것이다. 중요한 것은 이 글이 무료 게임과 온라인 게임에 초점이 맞추어져 있지 않다는 점이다. 즉 분석학은 모든 게임에 유용하다.

분석 대상 데이터

게임 분석을 위한 세 가지 주요 데이터 원천은 다음과 같다.

퍼포먼스(performance) 데이터: 이것은 게임의 기술적인 기반과 소프트웨어 기반의 실행과 관련된 데이터로, 특히 온라인 게임이나 지속적 게임(persistent game)과 관계가 있다. 퍼포먼스 매트릭스는 일반적으로 게임이 클라이언트 하드웨어 플랫폼에서 실행되는 프레임레이트 또는 게임 서버의 안정성을 포함한다.

프로세스 데이터: 이것은 게임을 개발하는 실제 과정과 관련된 것이다. 게임 개발은 어느 정도 창조적인 과정이지만, 예를 들어 업무 규모를 평가하는 것과 번다운 차트(burndown chart)를 사용하는 것을 통한 모니터링이 여전히 필요하다.

사용자 데이터: 가장 일반적인 데이터 원천인 이것은 개발자의 게임을 플레이하는 사용자들에게서 얻어진다. 개발자들은 사용자를 소비자(수익의 원천)로 봄과 동시에 *플레이어*, 즉 게임과 상호작용하게 될 때 특정한 방식으로 행동하는 사람으로도 본다. 첫 번째 관점은 사용자당 평균 매출(average revenue per user, ARPU)과 일간 접속 사용자(daily active users, DAU) 같은 수익과 관련된 매트릭스를 계산할 때, 또는 수익과 관련된 분석(사용자 이탈(churn) 분석, 소비자 지원 실행 분석, 또는 소액결제(microtransaction) 분석)을 수행할 때 사용된다.

두 번째 관점은 게임 내 행동(평균 플레이 시간, 세션당 받는 데미지 등)에 초점을 맞추므로써, 사람들이 실제 게임 시스템과 그 요소들 그리고 다른 플레이어들과 상호작용하는 방식을 조사하는 데 사용된다. 이것이 우리가 이 글에서 초점을 맞추게 될 데이터 유형이다. 이 세 가지 유형은 일반적인 **비즈니스 데이터**, 예를 들어 기업 가치와 기업 매출 등을 포괄하지 않는다. 우리는 이런 데이터를 게임

분석의 세부 영역이 아니라, 비즈니스 분석의 일반적인 영역에 속하는 것으로 간주한다.

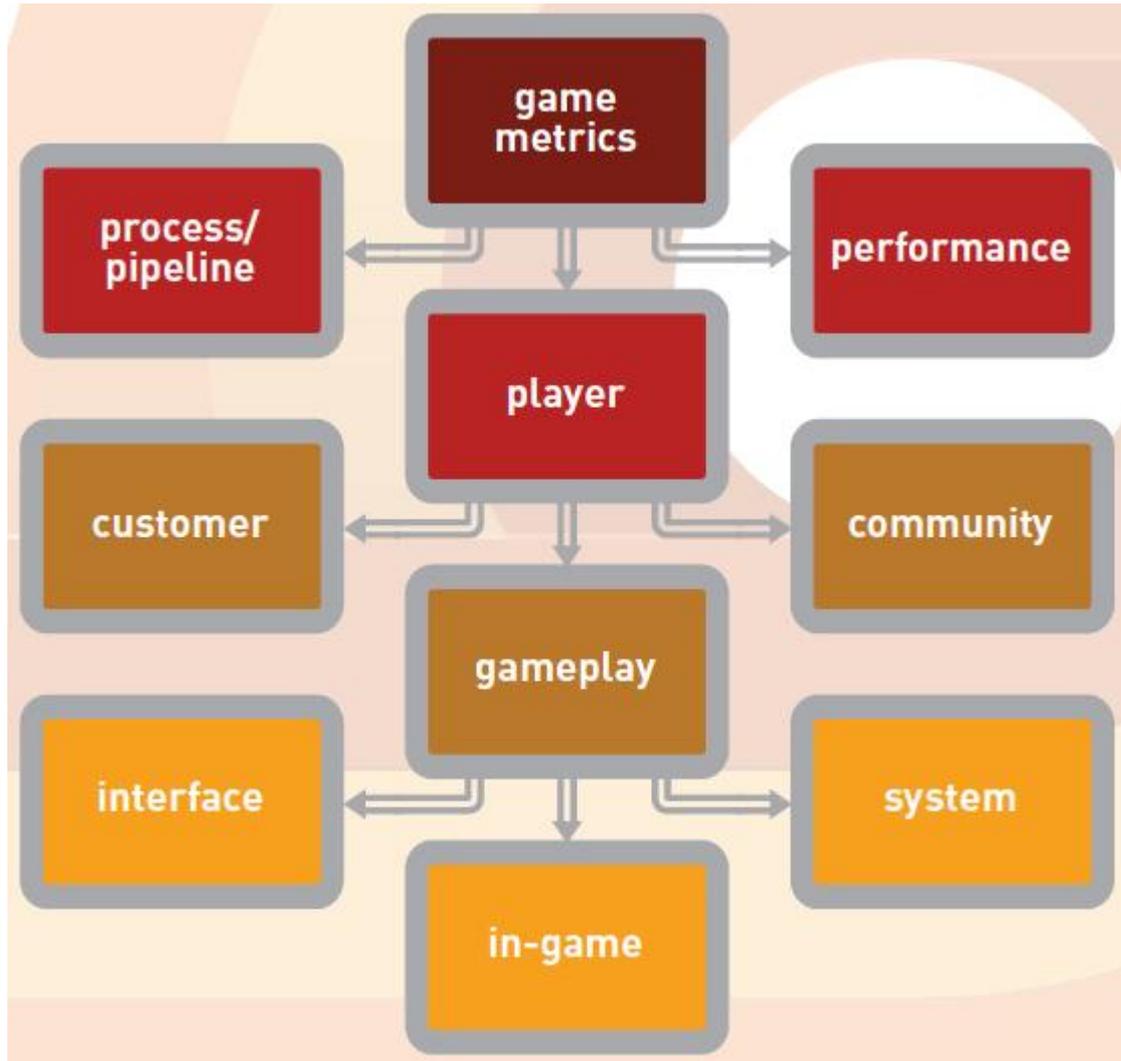


도표 1: 사용자 매트릭스를 강조하는 게임 분석을 위한 데이터 원천들의 위계적 다이어그램

사용자 데이터로 매트릭스 구성하기

지난 몇 년 동안 많은 사람들이 사용자 데이터를 분류하는 다양한 방법들을 제안했다. 하향적인 관점에서 보면 개발 중심의 분류 시스템이 유용하다. 이 시스템은 사용자 매트릭스를 세 가지 관련요소들(stakeholders)로 집중하는 데 도움이 되기 때문이다. 그 관련 요소들은 다음과 같다.

- **소비자 매트릭스:** 소비자로서 사용자의 모든 측면들, 예를 들어 소비자를 확보하고 유지하는 일의 비용 같은 것을 포괄한다. 게임과 게임 개발의 마케팅과 운영을 담당하는 전문가들에게 특히 유용하다.
- **커뮤니티 매트릭스:** 포럼 활동 같은 모든 수준의 사용자 커뮤니티 활동들을 포괄한다. 커뮤니티 운영자들에게 유용하다.
- **게임플레이 매트릭스:** 게임 내부의 플레이어로서 사용자가 취하는 실제 행동(예를 들어 object interaction, 물건 거래, 환경 속에서의 탐사)과 관련된 모든 변수를 말한다. 게임플레이 매트릭스는 게임 디자인과 사용자 경험을 평가하는 데 있어 가장 중요하며, 게임 개발에서의 수익 사슬에 대한 전통적인 관점과는 가장 관련이 없는 것이므로 보통 우선순위 아래에 있다. 디자인, 사용자 조사, 품질보증, 사용자의 실제 행동이 관심사가 되는 그 밖의 역할을 담당하는 전문가들에게 유용하다.

소비자 매트릭스: 소비자로서 사용자는 게임을 다운받아 설치할 수 있고, 짧거나 긴 시간 동안 실제 화폐나 게임 머니로 게임 내부나 외부의 상점에서 게임 아이템을 구매할 수 있다. 동시에 소비자는 버그를 보고하고 도움을 요청하고 불만 사항을 제출하는 등, 소비자 서비스와 상호작용한다. 또한 사용자는 공식적이거나 비공식적인 포럼들이나 소셜 네트워크 플랫폼들과도 상호작용하는데, 여기서 사용자들의 행동, 게임에 대한 만족도에 대한 정보를 얻고 분석할 수 있다. 개발자는 국적과 IP 주소, 때로는 나이와 성별, 이메일 주소 같은 소비자에 대한 정보도 수집할 수 있다. 이런 인구 통계학적인 정보와 행동에 대한 데이터를 결합하면 게임의 소비자 기반에 대한 강력한 통찰을 줄 수 있다.

커뮤니티 매트릭스: 사용자들은 기회가 있으면 서로 상호작용한다. 이 상호작용은 게임플레이(전투나 게임 메카닉을 통한 협력)나 소셜 네트워크(게임 내 채팅)와 관련될 수 있다. 플레이어들 간의 상호작용은 게임 내부나 외부에서, 또는 그것들의 결합, 예를 들어 페이스북에 올리기 기능을 사용해 새로운 장비를 획득한 것을 자랑하는 메시지를 보내는 것을 통해 이루어질 수 있다. 사용자들은 게임 내적으로는 채팅 기능을 통해, 외적으로는 실시간 대화(TeamSpeak 나 스카이프)나 포럼을 통해 다른 사용자들과 상호작용한다.

플레이어들 간의 이런 상호작용은 일련의 맥락들에서 사용할 수 있는 중요한 정보의 원천이 된다. 예를 들어 무료게임에서의 사용자 커뮤니티에 대한 소셜 네트워크 분석은 강력한 소셜 네트워크를 가진 플레이어들을 보여준다. 이들은

좋은 소셜 네트워크 환경을 만들어냄으로써 많은 수의 플레이어들을 게임에 남아있도록 하는 데 도움을 줄 수도 있다(MMORPG 게임의 길드 리더들을 생각해 보라). 마찬가지로 채팅 로그와 포럼 글들을 수집하는 것은 게임의 디자인에서 발생한 문제점들에 대한 정보를 줄 수 있다. 예를 들어 온라인 게임의 채팅 로그에서 얻은 데이터를 모으는 데이터세트들은 버그들이나 다른 문제점들을 보여준다. 플레이어들 간의 상호작용을 모니터하고 분석하는 것은 다수의 플레이어가 존재하는 모든 상황에서 중요하지만, 지속적인 플레이어 커뮤니티를 만들고 지원하고자 하는 게임들과 많은 소셜 온라인 게임과 무료게임을 포함하는 온라인 사업 모델을 적용하는 게임의 경우에 특히 중요하다. 이 사례들은 빙산의 일각일 뿐이며, 플레이어들 간의 상호작용에서 얻은 게임 매트릭스에 대한 수집과 분석, 보고는 책 몇 권은 나올 수 있는 주제다.

게임플레이 매트릭스: 이 사용자 매트릭스는 현재 쓰이고 있는 게임 원격 측정법 중 가장 널리 기록되고 사용되는 유형일 것이다. 게임플레이 매트릭스는 탐사, 아이템과 능력 사용, 점프, 거래, 달리기, 그 밖에 플레이어들이 게임의 가상 공간(2D 든 3D 든)에서 실제로 하는 모든 행동들의 척도이다. 게임에서 플레이어가 어떤 행동을 하거나 플레이어에게 어떤 일이 일어날 때마다 네 가지 유형의 정보, 즉 '무슨 일이 일어나고 있는가', '어디서 일어나고 있는가', '언제 일어나고 있는가', '누가 관련되어 있는가'에 대한 정보가 기록될 수 있다.

게임플레이 매트릭스는 게임 디자인에 정보를 제공하는 데 특히 유용하다. 이것은 만약 플레이어가 게임의 특성들을 의도한대로 사용한다면 어떤 게임 세계가 너무 많이 사용되는지 아니면 충분히 이용되지 않는지, 그리고 플레이어의 발전에 방해하는 한계들이 있는지 등 핵심적인 질문들을 제시하는 기회를 준다. 이 게임 매트릭스는 게임 개발의 모든 과정이 진행되는 동안은 물론, 게임 출시 이후에도 기록될 수 있다.

플레이어는 하나의 게임 세션이 진행되는 동안 수천 가지의 행동의 척도들을 생산할 수 있다. 즉 플레이어가 게임 시스템에 어떤 것을 투입할 때마다 시스템은 반응해야 한다. 플레이어의 활동에 대한 정확한 측정은 초마다 측정된 아주 많은 행동들을 포함할 수 있다. 예를 들어 <월드 오브 워크래프트> 같은 전형적인 판타지 MMORPG 게임을 하고 있는 플레이어를 생각해 보라. 사용자의 행동을 측정하는 것은 플레이어의 캐릭터의 위치, 현재 건강, 마나, 체력, 어떤 버프가 그 캐릭터에게 영향을 주는 시간, 적극적인 행동(달리기, 무기 휘두르기), 모드(전투,

거래, 여행), 플레이어에 대한 어떤 NPC 적들의 태도, 플레이어 캐릭터의 이름, 종족, 레벨, 장비, 골드 등을 기록하는 것을 포함할 수 있다. 이 모든 정보들은 단순히 설치된 게임 클라이언트에서 collection servers 로 흘러간다.

실용적인 관점에서, 개발자들은 (예를 들어 매트릭스를 찾기 쉬운 것으로 만들기 위해) 게임플레이 매트릭스를 다음의 세 가지로 더 세분화하기를 원할 수도 있다.

- **게임 내부:** 탐사와 경제 활동, 사물과 생명체 같은 게임 자산들과 상호작용 등 플레이어가 게임 안에서 하는 모든 활동들을 포괄한다. 이것은 대부분의 경우 수집된 사용자의 원격 정보가 된다.
- **인터페이스:** 플레이어가 게임 인터페이스와 메뉴와 행하는 모든 상호작용들을 포함한다. 이것은 마우스 감도와 화면 밝기 같은 게임 변수들을 설정하는 것을 포함한다.
- **시스템:** 시스템 매트릭스는 게임 엔진들과 그 하부 시스템들(인공지능 시스템, 자동화된 사건들, MOB/NPC 의 활동들 등)이 플레이어의 활동에 반응하도록 되어 있는 활동들을 포괄한다. 예를 들어 공격 범위 안에 들어온 플레이어 캐릭터를 공격하는 MOB, 또는 미리 정해진 조건들을 만족시킨 플레이어를 다음 레벨로 진급시키는 것이 있다.

요약하자면, 게임(또는 게임 서비스) 사용자들로부터 얻을 수 있는 일련의 잠재적인 척도들은 놀랄 만큼 많으며, 일반적으로 개발자들은 가장 중요한 정보를 기록하고 분석하는 것을 목표로 삼아야 한다. 이 선택 과정에서 편향성이 생길 수 있지만, 데이터가 지나치게 많아지는 것을 피하고 분석에서 기능적인 작업의 흐름을 확실하게 하는 데 필수적인 경우가 많다.

분석의 통합

모니터하기 위한 특징을 선택하고, 적용된 전략들을 측정하는 데 데이터 세트에 편향성이 들어가며, 이런 일은 분석가들이 고립된 상태에서 작업할 때 많이 발생한다. 분석에 책임이 있는 사람들이 관련된 모든 의사결정권자들과 소통할 수 없다면, 결정적인 정보가 사라질 것이 분명하고 분석의 가치가 완전히 실현되지 않을 것이다.

분석 집단들은 회사마다 다른 위치에 처하게 되는데, 이것은 분석학이 특히 사용자 조사, 마케팅, 수익화 같은 다양한 방향에서 산업에 도입되기 때문이다. 이런 현상은 분석 팀이 자신의 상부를 위해서만 일하고 상부만 우선시하는 상황으로 이어질 수 있다. 강력한 횡적인 통합, 예를 들어 분석 팀이 모든 팀과 소통하도록 하는 것은 이런 문제점을 피하는 데 도움이 된다. 또한 이렇게 하면 분석팀이 기획팀에 충분히 접근할 권한이 없을 때, 게임 디자인과 수익화 모델에 대한 적절한 기반 없이 추적하고 분석할 특징들을 선택할 수 밖에 없게 되는 일반적인 문제점을 완화하는 데에도 도움이 된다.

소규모의 개발사가 파트타임 분석가와 일하는 경우에도 이것이 문제가 될 수 있다. 또 다른 전형적인 문제점은 추적할 행동들을 결정하는 것이 분석 팀의 참여 없이 이루어진다는 점이다. 이것은 나중에 꼭 필요하지 않은 데이터나, 데이터세트를 추가로 기록해야 하는 데이터를 가지고 작업하느라고 많은 시간을 소비하는 결과로 이어질 수 있다. 여러 팀 간의 원활한 소통은 분석과 디자인 간의 충돌을 완화하는 데 도움이 된다.

중요한 것은 분석이 제작 시작 단계부터, 즉 초기 디자인 단계들로 거슬러 올라가 통합되어야 한다는 것이다. 어떤 종류의 행동을 어떤 빈도로 추적할 것인지를 초기에 계획해야 한다. 이것은 분석에서 디자인, 수익화, 마케팅 등에 이르는 평가를 어떻게 확실하게 할 것인지에 대한 최선의 계획을 세울 수 있게 해준다. 분석은 베타 테스트 이후에 이루어져서는 안 된다. 이런 방식으로 분석은 개발 과정 내내 그리고 출시 이후에도 이상적으로 개입된다는 점에서 사용자 조사 같은 다른 도구들과 비슷하다.

특징 선택

사용자의 행동에 대해 개발자들이 측정할 수 있는 일련의 데이터가 있다는 것을 알고 난 다음에는, 어떻게 그것들 사이에서 필요한 데이터를 선택할 것인가? 정말로 여기서 선택을 해야 하는가? 슬프지만 그렇다. 현실적으로 가능한 모든 사용자의 행동들을 추적하고 분석하기 위한 자원이 별로 없다. 따라서 한편으로는 사용자의 원경 정보/매트릭스를 추적하고 저장하고 분석하는 데 필요한 자원들 간의 비용-편익 관계를, 다른 한편으로는 획득한 통찰의 가치를 고려해 접근법을 개발해야 한다는 것을 의미한다. 다양한 제작과 출시 이후 단계에서 각기 다른

분석이 필요하다는 점을 아는 것도 중요하다. 예를 들어 개발 후기 단계에서는 디자인을 조절하는 것이 중요하지만, 목표 소비자들이 아직 게임을 구매하지 않았으므로 수익화와 관련된 많은 매트릭스는 측정할 수 없다.

우리는 이것을 좀더 상세하게 논할 것이다. 그런데 요약하자면, 이 논리를 따름으로써 추적하고 저장하고 분석해야 하는 사용자 특징은 최소한 다음에 대해 고려해야 한다.

1) 일반 특징: 모든 게임에서 (소비자이자 플레이어로서의) 사용자들에게 공유되는 특징들을 말한다. 이것은 모든 컴퓨터 게임에서 항상 수집될 수 있는 핵심 매트릭스가 된다. 예를 들어 사용자가 게임을 시작하거나 중단하는 시간, 사용자 ID, 사용자 IP, 진입점(entry point) 등이 있다. 이것은 모든 게임 분석 데이터세트의 핵심이 된다.

2) 핵심 메카닉/디자인 특징: 게임플레이와 게임 메카닉의 핵심과 관련된 기본적인 특징들(예를 들어 플레이에 소비된 시간, 사이버 머니 소비, 물리친 적의 수 등과 관련된 특징들)을 말한다. 핵심 디자인 특징을 정의하는 것은 게임의 핵심 게임플레이 메카닉에 직접적으로 기초해야 하며, 디자이너들이 사용자 경험에 대한 추론(만약 게임플레이가 지속적으로 흐른다면 플레이어는 계획대로 진급할 것인지, 죽는 비율, 레벨 완료, 점수)을 할 수 있게 해주는 정보를 주어야 한다.

3) 핵심 비즈니스 특징: 기업의 비즈니스 모델의 핵심과 관련된 기본적인 특징들로, 예를 들어 사용자가 매번 사이버 아이템을 구매하고(그리고 그 아이템이 무엇인지), 게임 안에서 친구 관계를 형성하거나 페이스북 친구에게 게임을 추천할 때, 또는 수익, 사용자 유지, 유행성, 사용자 이탈과 관련된 그 밖의 특징들을 말한다. 모바일 게임에서는 지리적 위치 데이터가 타겟 마케팅(target marketing)을 보조하는 데 아주 유익할 수 있다. 물론 이 특징들은 전통적인 소매시장에서는 유익하지 않다

4) 이해관계자 요구사항: 여기에 더해 다양한 '이해관계자 요구사항들(stakeholder requirements)'도 고려해야 한다. 예를 들어 일간 접속 사용자의 수를 알아내는 것에 높은 가치를 두는 경영진이나 마케팅팀 말이다. 이런 요구사항들은 위에서 언급한 범주들과 연결될 수도 있고 그렇지 않을 수도 있다

5) 품질 보증과 사용자 조사: 마지막으로, 만약 사용자 조사/사용자 테스트와 품질 보증을 위해 원격 정보를 사용하는 일(충돌과 그 원인, 클라이언트 시스템의 하드웨어 배치, 중요한 게임 설정을 기록하는 것)이 조금이라도 유익하다면, 특징 목록의 특징들을 적절히 증가시켜야 할 수도 있다.

초기 특징의 집합을 구성하고 그로부터 획득할 수 있는 매트릭스를 계획할 때에는, 선택 과정이 정보를 최대한 잘 주고 모든 관련요소를 포함하는지 확인해야 한다. 그래야 나중에 코드로 돌아가 추가로 작성해 넣어야 하는 일이 줄어든다. 즉 이런 낭비는 세심한 계획이 있다면 제거할 수 있다.

말하자면 게임은 (지속적인 게임이든 DLC/패치를 통해서든) 제작 기간 동안에는 물론 출시 이후에도 발전하기 때문에, 새로운 특징들을 추적하고 발전하는 분석을 지속적으로 실행하기 위해 코드에 새로운 요소를 추가해야 한다. 샘플을 만드는 것도 고려해야 할 중요 사항이다. 누군가가 총을 쓰는 모든 순간을 추적할 필요는 없고 그 중 1%만 추적하면 된다. 샘플을 만드는 것은 그 자체만으로도 큰 사안이므로, 이 글에서는 그것이 게임 분석을 위해 필요한 자원들을 골라내는 효율적인 방법이 될 수 있다는 것만 지적하고 더 이상의 논의는 진행하지 않을 것이다.

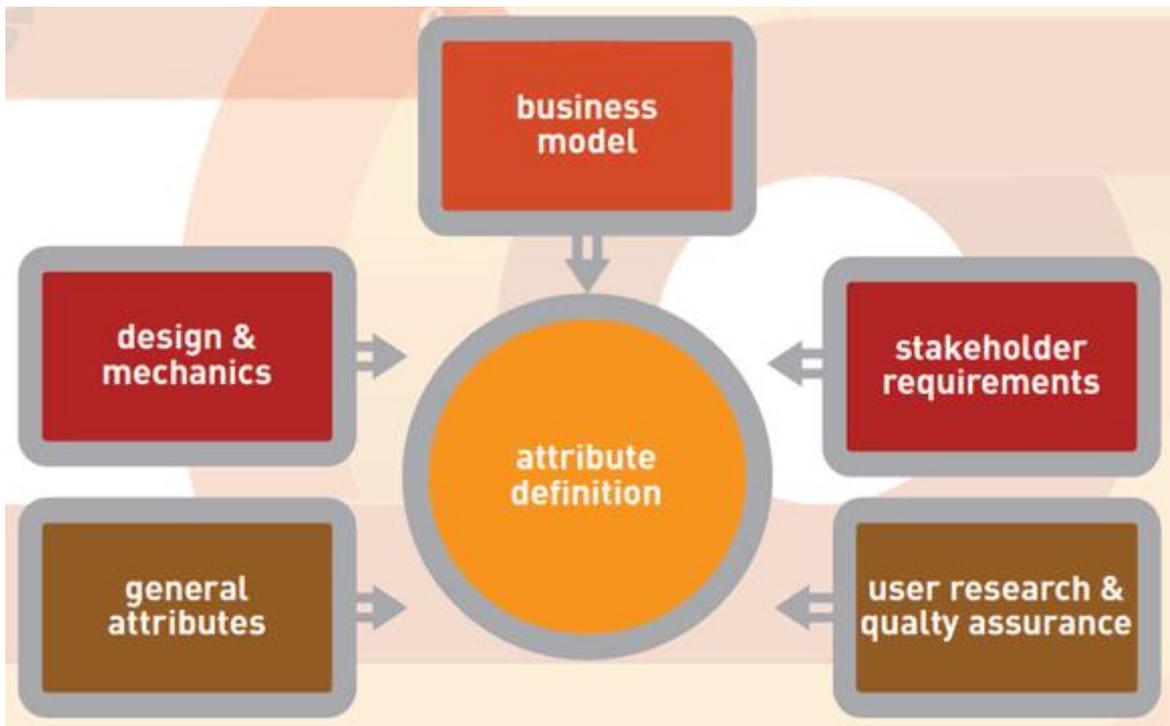


도표 2: 사용자 행동 특징 선택의 드라이버. 게임 분석의 폭넓은 적용 범위를 고려해, 몇 가지 필요한 자원들이 존재한다.

특징의 사전 선택

특징 선택 과정에서 고려해야 할 중요한 한가지 요소는 특징들 선택하는 과정이 사전 계획에 의해, 그리고 개발자가 사용자 원격 정보에서 획득하고 그에 따라 특징들을 선택하고자 하는 게임 매트릭스와 분석 결과들(그리고 그것에 의해 실행 가능한 지식들)을 정의하는 것에 의해 진행될 수 있다는 점이다.

복잡성을 줄이는 것은 필수적이지만, 데이터 수집 과정의 범위를 제한했기 때문에 개발자들은 사전에 선택된 특징들을 사용하면서 찾아낼 수 없는 사용자 행동의 중요한 패턴들을 놓치는 위험을 안게 된다. 이 문제점은 게임 매트릭스와 분석이 사전에 정의된 상황에서 악화된다. 예를 들어 (일간 접속 사용자, 월간 접속 사용자(MAU), 사용자당 평균 매출, LTV 등과 같은) 일련의 핵심성과지표들(Key Performance Indicators)에 의존하는 것은 사전에 정의된 매트릭스와 분석으로는 찾아낼 수 없는 사용자 행동의 패턴들을 발견할 수 있는 기회를 없애버릴 수 있다.

일반적으로는 사용할 수 있는 분석 자원에 의존해 두 상황들 간의 균형을 맞추는 것이 최선의 해결책이다. 예를 들어 핵심성과지표들에만 집중하는 것은 35%의 플레이어들이 레벨 8 에서 포기하는 이유 같은 게임 내부의 행동에 대한 정보를 주지 않는다. 따라서 개발자들은 디자인과 실행과 관련된 매트릭스를 조사해야 한다.

사용자 분석을 하게 될 때 개발자들은 예측할 수 없는 인간의 행동을 대상으로 작업한다는 것을 지적할 필요가 있다. 이것은 사용자 분석의 요구사항들을 예측하는 것이 어려울 수 있다는 것을 의미한다. 이것은 예비적인 방법(개발자들은 그것이 어떤 패턴들을 포함하는지 알아내기 위해 사용자 데이터를 조사한다)과 가설에 의해 진행되는 방법(개발자들은 자신이 측정하고자 하는 것이 무엇인지 알고 있으며 어떤 것이 정확한 것인지 뿐만 아니라 가능한 결과들도 알고 있다) 모두를 사용해야 할 필요성을 강조한다.

기획자의 지식에 의해 진행되는 전략들

게임을 플레이하는 동안 사용자는 게임 상태의 변화를 이어가는 행동과 반응을 끊임없이 반복한다. 이것은 어떤 순간에서든 가치를 변화시키는 사용자 행동의 특징들이 많다는 것을 의미한다. 분석 과정에서 어떤 특징들을 사용할 것인지를 선별하는 일의 첫 단계는 게임과 플레이어들 사이에서 일어날 수 있는 모든 상호작용들의 종합적이고 상세한 목록일 것이다. 디자이너들은 게임과 플레이어들 사이에서 일어날 수 있는 모든 상호작용들에 대한 지식을 아주 많이 가지고 있다. 그 지식을 사용하고 디자이너들에게 그런 목록들을 작성하도록 요청함으로써 시작 단계부터 그들과 함께 하는 것은 유익하다.

다음으로는, 가장 단순한 게임에서조차 게임에 포함되어 있는 변수들의 숫자를 고려하면서 지식에 의해 진행되는 요소를 줄임으로써 복잡성을 줄여야 한다. 디자이너들은 형태가 같은 상호작용들을 쉽게 식별할 수 있다. 이것들은 형태는 조금씩 다르지만 본질적으로는 비슷한 상호작용들, 행동들, 상태 변화들의 집합들이다. 예를 들어 "붕대로 5 HP 회복"과 "물약으로 50 HP 치료"는 형태는 달라도 본질적으로는 비슷한 행동이다. 그런 다음 형태가 같은 상호작용들은 더 넓은 영역으로 묶여진다. 마지막으로, 각각의 영역에 속하는 형태가 같은 모든 상호작용들을 포착하는 척도들을 식별해야 한다. 예를 들어 "치료"의 영역에서는

사용된 물약과 봉대의 수를 추적할 필요가 없지만, "건강"이라는 변수에 발생하는 모든 상태 변화는 기록해야 한다.

이 영역들은 객관적인 요소를 줄임으로써 얻어지는 것이 아니다. 디자이너가 광범위한 전문지식을 가지고 있다고 해도, 요소들을 범주로 묶어야 할 때에는 언제나 해석상의 분명한 편향성이 존재한다. 더 넓은 이 영역들은 플레이어들이 게임 내에서 할 수 있는 모든 행동들을 잠재적으로 포함할 수 있고 어떤 게임 변수들을 어떻게 모니터해야 하는지를 선택하는 데 도움을 준다.

기계 학습에 의해 진행되는 전략들

기계 학습(machine learning)은 명시적으로 프로그램되지 않고도 학습할 수 있는 능력을 컴퓨터에 부여하는 학문 분야이다. 디자이너에 의해 진행되는 전략에 대한 대안 이상으로, 자동화된 특징 선택은 플레이어와 게임 사이의 상호작용들에 의해 발생하는 수많은 상태 변화들의 복잡성을 줄이는 것에 대한 보완적인 접근법이다. 전통적으로 자동화된 접근법은 존재하는 데이터세트와 관계형 데이터베이스(relational database) 또는 데이터 웨어하우스(data warehouse)에 사용되는데, 이것은 게임 시스템을 분석하고 변수들을 정의하고 그 변수들의 척도들을 확립하는 과정이 자동화된 전략의 범위에서 벗어난다는 것을 의미한다. 즉 인간은 어떤 변수들을 어떻게 추적할 것인지를 미리 정의한다. 따라서 자동화된 접근법은 모니터되는 모든 변수들 중에서 관련성이 가장 높고 가장 특이한 특징들만 선별해낸다.

자동화된 특징 선택은 다른 것들과 관련성이 높은 특징 공간과 drop feature 들을 찾아내기 위한 알고리즘에 의존한다. 이 알고리즘들은 간단한 것에서 복잡한 것까지 이를 수 있다. 방법은 클러스터링, 분류, 예측, 순차 데이터 처리 같은 접근법을 포함한다. 이 접근법들은 관련성이 가장 높은 특징들을 찾아내는 데 사용될 수 있는데, 왜냐하면 유형을 정의하는 것과 관련이 없는 특징들의 존재는 알고리즘으로 찾아낸 클러스터들의 질을 떨어트리면서 유사성 측정(similarity measure)에 영향을 주기 때문이다.

수확 체감

자원이 무한한 상황에서는 모든 서버-사이드 시스템 정보, 아바타의 모든 동작, 모든 구매 내역, 채팅 메시지, 버튼 터치, 심지어 키보드 터치 같은 사용자로부터 시작되는 모든 활동들을 추적하고 저장하고 분석할 수 있다. 이렇게 하는 것은 대역폭 문제를 일으킬 수 있으며, 게임 코드에 메시지 요소들을 추가하기 위해 상당한 자원을 필요로 할 것이다. 이론적으로는 이런 역지스러운 방법으로 게임 분석에 접근하는 것이 가능하다.

하지만 이런 접근법은 엄청난 양의 데이터세트로 이어져 결국 그것들을 변형하고 분석하기 위해 엄청난 양의 자원을 필요로 하게 된다. 예를 들어 무기의 유형, 무기 개조, 범위, 데미지, 표적, 죽인 횟수, 플레이어와 표적의 위치, 총알의 궤적 등을 추적하는 것은 FPS 게임에서의 무기 사용에 대한 깊이 있는 분석을 할 수 있게 해줄 것이다. 그러나 무기의 균형을 측정하는 핵심 매트릭스는 범위, 데미지, 무기들 각각의 사용 빈도에 지나지 않을 수도 있다. 변수/특징 몇 가지를 추가하는 것은 새로운 관련 지식들을 주지 않을 수도 있고, 오히려 분석에 혼란을 초래할 수도 있다. 마찬가지로 모든 플레이어의 행동의 원격 정보를 기록할 필요는 없고, 백분율만 기록하면 될 것이다(물론 매출을 기록하는 경우에는 다른데, 개발자들은 모든 매출을 추적해야 하기 때문이다).

일반적으로, 변수/특징들을 제대로 선택했다면 추적하고 수집하고 분석한 첫 번째 변수들/특징들은 사용자의 행동에 대한 많은 지식을 줄 것이다. 사용자 행동의 상세한 양상들을 더 많이 추적할수록 저장과 처리, 분석에 들어가는 비용은 증가하지만, 원격 측정 데이터에 포함된 정보로부터 추가된 가치의 비율은 감소한다.

이것은 게임의 원격 측정에 비용-편익 관계, 즉 단순화한 '수확 체감의 법칙'으로 표현되는 것이 있다는 것을 의미한다. 분석 과정에서 하나의 데이터 자원의 양을 증가시키면 단위당 산출량이 낮아진다는 것이다.

경제학 문헌에서는 토양에 비료를 뿌린다는 고전적인 사례를 들고 있다. 균형이 맞지 않는 시스템(비옥하지 않은 땅)에 비료를 뿌리는 것은 수확량을 증가시키겠지만, 어느 시점 이후에는 이 증가세가 줄어들고 멈추며, 심지어는 수확량을 감소시킬 수도 있다. 이미 균형이 맞춰진 시스템에 비료를 뿌리는 것은 수확량을 증가시키지 않거나 감소시킬 수도 있다.

기본적으로 게임 분석은 비슷한 법칙을 따른다. 분석은 추가적인 (새로운) 특성들이 필요해지기 전에 특징들/변수들의 특정한 집합들이 투입되는 특정한 시점까지 최적화될 수 있다. 게다가 분석 과정에 투입되는 데이터의 양을 늘리는 것은 산출량을 감소시킬 수도 있고, 아니면 극단적인 경우 데이터를 추가함으로써 생긴 혼란으로 인해 부정적인 산출이 나오는 상황으로 이어질 수 있다. 물론 예외는 있을 수 있다. 예를 들어 소셜 온라인 게임에서 사용자 보유를 감소시키는 문제 있는 행동 패턴의 원인은, 그 결함과 관련된 특정한 행동 변수들을 추적하지 않는다면 확인하기 어려울 수 있는 디자인상의 작은 결함 하나에 있을 수 있다.

사용자 중심 분석의 목표

사용자 중심의 게임 분석에는 보통 다양한 목표들이 있지만, 아래와 같이 광범위하게 구분할 수 있다.

- **전략적 분석:** 사용자의 행동과 비즈니스 모델에 대한 분석에 기초해 게임이 어떻게 발전할 것인지에 대한 전체적인 고찰을 목표로 한다.
- **전술적 분석:** 단기간의 게임 디자인에 정보를 제공하는 것을 목표로 한다. 예를 들어 새로운 게임 특징에 대한 A/B 테스트가 있다.
- **실행 분석(operational analytics):** 게임에서의 즉각적이고 진행중인 상황에 대한 분석과 평가를 목표로 한다. 예를 들어 실시간으로 사용자 행동을 어울리도록 하기 위해 지속적인 게임에 어떤 변화를 주어야 하는지에 대한 정보를 제공하는 것이다.

어느 정도는, 실행 분석과 전술적 분석은 기술과 기초상의 사안들에 정보를 제공하고, 전략적 분석은 사용자의 원격 측정 데이터를 다른 사용자 데이터 및 시장 조사와 통합하는 것에 초점을 맞추고 있다.

사용자의 원격 측정에 접근하기 위한 전략을 짤 때 신경 써야 할 첫 번째 요소들은 사용자 중심의 게임 분석에는 이 세 가지의 유형이 존재한다는 것과 이 유형들이 요구하는 입력 데이터의 종류들, 이 유형들이 모두 실행되고 있다는 것을 확실하게 하기 위해 해야 하는 것, 관련요소에 보고된 결과 데이터이다.

두 번째로 고려해야 할 요소는 어떻게 기업의 욕구와 사용자의 욕구 모두를 만족시킬 것인지를 분명하게 하는 것이다. 게임 디자인의 근본적인 목표는 사용자에게 좋은 경험을 제공하는 게임을 만드는 것이다. 하지만 게임 개발 기업을 운영하는 근본적인 목표는 (적어도 투자자들의 관점에서는) 돈을 버는 것이다. 분석 과정이 이 두 가지 목표 모두를 달성하기 위한 의사결정을 뒷받침하는 결과를 산출한다는 것을 확실하게 하는 것이 중요하다. 기본적으로 게임 분석의 기초 드라이버는 다음의 두 가지 요소로 이루어진다. 1) 소비자를 획득하고 유지하기 위해 양질의 사용자 경험을 보장하는 것, 2) 논의되고 있는 비즈니스 모델과 무관하게 매출을 올리는 수익화 사이클을 보장하는 것이 그것이다. 사용자 중심의 게임 분석은 디자인과 수익화 모두에 동시에 정보를 제공해야 한다. 이 접근법은 디자인상의 특정한 변화가 사용자의 경험(때로는 사용자 보유가 그것을 평가하기 위한 대용으로 사용된다)과 수익을 증가시키는지를 평가하기 위해 A/B 테스트 같은 분석 방법을 사용해 무료게임 시장에서 성공한 기업들의 사례로 증명되었다.

요약

지금까지 특징 선택에 대한 논의를 구체적인 매트릭스(분당 무기 발사 비율, 죽은 횟수와 죽은 횟수의 비율, 점프 성공률)의 목록을 작성하기보다는, 다소 추상적인 수준에서 선택의 가이드가 되는 범주를 만들고 포괄적으로 다뤄왔다. 이것은 모든 타입의 게임과 사용 상황들을 아우르는 매트릭스의 일반적인 가이드라인을 개발하기가 거의 불가능하기 때문이다. 게임들이 정해진 디자인 유형에 들어맞지 않기 때문일 뿐만 아니라(게임들은 광범위한 디자인 영역을 공유하며 그 중 특정한 영역들로 몰리지 않는다), 디자인상의 혁신 비율이 높기 때문이기도 하다. 여기서 후자는 추천하는 것을 급속하게 부당하게 만든다. 따라서 사용자 분석에 대해 우리가 해줄 수 있는 조언은 위로부터는 모델들을 개발함으로써 데이터 수집에 있어 포괄적인 적용범위를 확보할 수 있도록 하고, 중심으로부터는 (디자이너들을 돕기 위해서는) 사용자의 경험과 (디자이너들이 급여를 받는 것을 확실하게 하는 데 도움을 주기 위해서는) 수익화를 운영하는 주요 메카닉에서 시작하라는 것이다. 자원이 허락한다면 추가적인 자세한 사항이 추가될 수 있다. 마지막으로, 자신의 결정과 과정을 유동적이고 변형 가능하게 유지하도록 노력하라. 게임 산업만큼 경쟁적이고 흥미로운 산업에서 이것은 필수적이다.

[Magy Seif El-Nasr, Anders Drachen, Alessandro Canossa 는 최근에 출간된 50 명 이상의 게임 산업 전문가와 조사 전문가들의 지식들을 요약한 책 <Game Analytics: Maximizing the Value of Player Data>²의 편집자이다. 이 글은 그 책에서 발췌한 글들을 바탕으로 쓰여졌다.]

² 참조링크: <http://link.springer.com/book/10.1007/978-1-4471-4769-5/page/1>