데이터의 수집: 경로와 현황

I. 서론

II. 행태정보 수집 방법론

1. 행태정보 수집 개요
2. 쿠키의 개념과 기능
3. 광고 식별자의 개념과 기능 : AdID, IDFA 등
4. 소결

III. 국내 데이터 수집 현황

1. 쿠키를 이용한 데이터 수집
2. AdID를 이용한 데이터 수집

IV. 온라인 광고 생태계의 구조

1. 온라인 광고 시장 개관
2. 온라인 광고의 유형
3. 온라인 광고 시장의 주요 행위자

V. 법적 현안 - 결론을 대신하여

고학수
서울대학교 법학전문대
학원 교수

김종윤
서울대학교 법학대학원
박사과정·변호사

김병필
KAIST 기술경영학부
초빙교수·변호사

I. 서론

많은 사람들에게 인터넷은 일상 생활의 중요한 일부이다. 여러 웹사이트를 방문하거나 스마트폰의 앱을 이용하여 다양한 업무를 처리하는 것은 더 이상 새로운 일이 아니다. 그러나 대부분의 인터넷 이용자들은 자신들이 이용하는 서비스의 뒷면에서 어떠한 데이터가 오고 가는지 구체적으로 알지 못한다. 인터넷 공간에서는 서비스에 필요한 정보뿐만 아니라, 이용자에 관한 행태정보가 전송되기도 한다. 이용자 행태 정보는 이용자가 이름을 들어 본 적도 없는 업체에 보내지기도 한다. 이용자가 이러한 사정을 모르는 것을 두고 개별 이용자의 관심 부족 때문이라고 치부하기도 어렵다. 이용자에 관한 정보가 수집되는 것과 관련한 구체적인 사항들이 잘 알려져 있지 않을 뿐만 아니라 그에 관한 사회적 논의도 부족하기 때문이다. 다른 한편, 이용자에 관한 여러 정보들은 맞춤형 서비스를 위해 이용되기도 하고, 온라인 광고 산업에 활용되어 인터넷 생태계를 지탱하는 역할을 하고 있기도 하다. 이 글은 인터넷 공간에서 데이터 수집이 이루어지는 기초적인 기술적 방법을 소개하고, 국내 인터넷 환경에서 데이터 수집이 실제로 어느 정도로 이루어지고 있는지에 관한 실증 조사 결과를 살펴본다. 나아가 인터넷 광고온라인 광고 생태계의 간략한 구조와 규모를 개관하고, 법적으로는 어떠한 현안들이 있는지 간략히 검토한다.

II. 행태정보 수집 방법론

1. 행태정보 수집 개요

PC와 스마트폰의 이용자들은 인터넷을 활용하는 양상이 다르다. PC에서는 대부분의 인터넷 이용이 웹브라우저를 통해 이루어진다. 이메일 클라이언트, 게임, 메신저, 몇몇 클라우드 앱 정도를 제외하고는 활용도가 높은 서비스의 상당 부분은 웹브라우저를 통해 제공된다. 반면 스마트폰을 통해 제공되는 온라인 서비스들은 앱을 통해 제공되는 경우가 많다. 유튜브 서비스를 예로 들면, PC 환경에서는 대다수 이용자가 웹브라우저를 통해서 서비스를 이용하지만, 스마트폰 환경에서는 유튜브 앱을 이용하는 경우가 많을 것으로 보인다.

이용자의 행태정보 수집은 이용자가 어떠한 방식으로 인터넷을 활용하는지에 따라 다른 방식에 의해 이루어진다. 이용자가 웹브라우저를 통하여 인터넷 서비스를 이용할 때에는 웹의 근간이 되는 HTTP(Hypertext Transfer Protocol)에 정의된 '쿠키(cookie)'를 이용하는 방법이 사용된다. 이에 비해, 다양한 앱이 사용되는 모바일 환경에서는 앱 사이의 공통적인 데이터 행태정보 수집 프레임이 필요하게 되는데, 아래에서 설명하는 AdID(Advertising ID)와 IDFA(Identifier for Advertisers) 등 광고 식별자가 그러한 공통 프레임의 역할을 담당해 오고 있

다. 이하에서는 (1) 웹브라우저상의 쿠키의 개념과 기능을 먼저 살펴보고, 다음으로 (2) 모바일 앱에서 이용되는 광고 식별자의 개념과 기능을 소개한다.

2. 쿠키의 개념과 기능

쿠키는 웹서버가 웹브라우저에 저장할 수 있는 작은 텍스트 파일이다. 이용자가 특정 웹사이트를 방문하는 과정은 기본적으로 해당 웹서버로부터 웹페이지의 내용을 불러와서 이용자의 웹브라우저를 통해 상호작용을 하는 것이라 할 수 있다. 이 과정에서, 웹서버는 이용자의 웹브라우저에 쿠키를 저장하도록 요청할 수도 있고, 웹브라우저에 저장되어 있던 쿠키를 조회할 수도 있다. 당초 웹 표준을 만드는 과정에서 쿠키를 고안하여 이용하게 된 이유는 웹사이트 이용자의 편의를 도모하기 위한 것이다. 예컨대, 이용자가 웹서버에 로그인하지 않은 상태에서 장바구니에 추가한 물건이 다음 방문 때에도 그대로 장바구니에 남아 있다면 이는 쿠키 덕분일 가능성이 높다. 웹서버는 이용자가 장바구니에 어떠한 물건을 추가해 두었는지에 관한 정보를 이용자의 웹브라우저에 쿠키 형식으로 저장해두었다가, 이용자가 다음번 방문할 때 저장된 쿠키에서 정보를 읽어와서 다시 장바구니에 추가할 수 있기 때문이다. 이용자가 한 차례 로그인을 하고 나면, 웹브라우저를 종료하였다가 다시 실행하여도 로그인 상태가 계속 유지되는 것 또한 쿠키를 이용하여 로그인 사실을 저장해 놓은 덕분일 수 있다.

인터넷 환경이 고도화되면서 '제3자 쿠키(third-party cookie)'라는 새로운 이용자 행태정보 수집 방식이 등장했다. 전통적인 쿠키는 이용자가 방문한 웹서버가 해당 이용자에 관한 정보를 저장하였다가 다시 불러오는 방식으로 동작한다. 이를 '당사자 쿠키(first-party cookie)'라고 한다. 이에 비해 제3자 쿠키는 이용자가 방문하고자 의도한 웹서버가 아닌 제3의 웹서버가 저장·조회하는 쿠키를 말한다. 인터넷의 전송 대역폭이 늘어나면서 웹페이지는 점차 복잡한 형태를 가지게 되었고, 하나의 웹페이지가 다양한 영역으로 분류되고 영역별로 다양한 정보가 채워지는 것이 일반화되었다. 예컨대, 인터넷을 통해 언론 기사를 접할 때, 이용자가 보게 되는 화면은 해당 기사 이외에도 실시간 인기 기사, 댓글, 광고 등 다양한 영역으로 나누어 볼 수 있다. 이러한 각각의 영역은 서로 다른 도메인의 여러 웹서버가 제공하는 독립적인 정보로 채워질 수 있다. 그 경우 해당 기사가 담긴 웹페이지를 방문하는 이용자의 입장에서는 마치 하나의 웹서버에 접속한 것처럼 이해할 수 있지만, 실제로는 여러 개의 서로 다른 도메인과 통신하고 있는 셈이다. 이때 당초 이용자가 접속한 웹서버 이외의 다른 도메인의 웹서버가 이용자 웹브라우저에 저장하는 쿠키가 제3자 쿠키이다.

제3자 쿠키는 이용자의 편의와 직접적인 관련이 없이 이용자에 관한 행태정보를 수집하는 용도로 활용되는 경우가 많다. 이를 통상 트래킹(tracking)이라 표현한다. 인터넷상에서 제3자 쿠키는 주로 온라인 광고업 등을 운영하는 회사들에 의하여 활용되는 것으로 보인다.[1] 쿠키 정보를 적극적으로 수집하는 기업

1) 고학수 외 2명, "국내 웹사이트의 이용자 정보수집 및 트래킹 현황에 관한 분석," 법경제학연구, 2017. 12., 428~431면

들에는 광고수익에 크게 의존하는 대형 온라인 플랫폼 사업자와 온라인 광고 전문사업자가 포함된다. 이들 기업이 정보를 수집하는 주된 목적은 이용자 트래킹을 통하여 맞춤형 광고를 집행하기 위한 것으로 보인다. 이는 국내는 물론 해외에서도 유사하다. 해외 연구의 한 예를 들면, 100만 개의 주요 웹사이트를 조사한 결과, 그중 10% 이상의 웹사이트에서 적극적으로 쿠키를 수집하는 제3자들은 모두 이용자에 대한 행태정보 수집, 즉 트래킹이 주목적이었다고 한다.[2)]

이용자 트래킹 사업자(즉, 제3자 서버 운영자)가 쿠키를 통해 이용자를 트래킹하는 가장 기본적인 방법은 다음과 같다. 트래킹 사업자가 이용자의 웹사이트 방문 이력 정보를 수집하고자 하는 상황을 생각해 보자. 트래킹 사업자는 다수의 웹사이트 제공자와 약정을 체결하고, 그 웹페이지의 한 영역에 트래킹 사업자의 웹페이지를 삽입한다. 트래킹 사업자가 삽입한 영역은 광고를 포함하고 있을 수도 있고, 아무런 내용이 표시되지 않을 수도 있다.[3)] 이용자가 웹브라우저를 이용하여 방문하고자 하는 웹페이지에 접속하면 트래킹 사업자의 웹서버에 있는 웹페이지도 함께 불러오게 된다. 이때 트래킹 사업자의 웹서버로 이용자가 어떤 웹페이지에 접속하였는지에 대한 정보가 전송된다. 이러한 방법을 통해 트래킹 사업자는 특정 이용자가 어떤 웹사이트에 접속했는지 확인할 수 있게 된다.

트래킹 사업자가 각기 다른 웹 서비스 제공자가 운영하는 A, B, C, D 웹사이트와 약정을 체결하고 이용자가 그 웹사이트 접속했는지에 관한 정보를 수집하는 상황을 생각해 보자(아래 그림 1 참조). 그림을 통해 예시된 상황의 경우에, 이용자가 웹사이트 A로부터 웹사이트 B, C, D 등으로 이동하더라도 트래킹 사업자(제3자 서버)의 웹서버와 이용자 사이의 연결은 계속해서 유지된다. 트래킹 사업자의 웹서버는 이용자를 트래킹하기 위해 이용자의 컴퓨터에 쿠키(즉 제3자 쿠키)를 저장한다. 이용자는 여러 인터넷 서비스 제공자의 웹사이트를 방문하지만, 동일한 트래킹 사업자의 서버와 제3자 쿠키를 통해 지속적으로 연결이 이루어지기 때문에 이용자의 웹페이지 접속 이력 등에 관한 정보가 트래킹될 수 있다. 즉, 쿠키 기능을 이용하면 이용자가 장바구니에 추가한 물건이 계속 남아 있는 것과 마찬가지로, 트래킹 사업자는 제3자 쿠키를 이용함으로써 이용자가 별도로

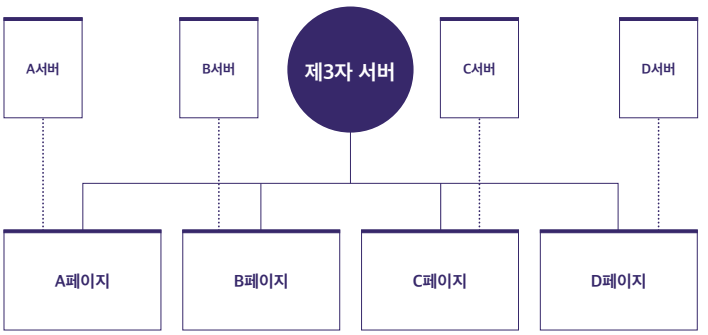

그림 1. 쿠키의 작동 방식[4)]

2)
Engelhardt, Steven et al., "Online Tracking: A 1-million-site Measurement and Analysis," Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security(CCS '16), 2016. 10., p. 1395

3)
이 경우, 소스코드에 대한 확인 없이 웹페이지만을 보아서는 이용자 입장에서는 제3자의 서버에 소재한 웹페이지가 로드되었는지조차 알 수 없다.

4)
각 웹페이지의 구획된 일부 공간은 별도의 웹서버에 소재해 있을 수 있으며, 그림 1의 경우 제3자의 웹서버에 소재한 것으로 표시되었다. 해당 일부 공간은 이용자가 볼 때에는 배너 광고일 수도 있고, 소셜 네트워크의 로그인 구역일 수도 있으며, 육안으로 식별할 수 없는 크기 및 색일 수도 있다.

트래킹 사업체의 웹서버에 로그인하지 않더라도 이용자의 여러 웹사이트 방문 이력을 파악할 수 있게 된다.

한편, 트래킹 사업자가 제3자 쿠키를 이용하여 이용자의 동일성을 식별하는 과정은 다음 그림 2를 통해 설명할 수 있다. 이용자가 처음으로 트래킹 사업자 웹서버에 접속할 때에는 이용자의 웹브라우저에는 제3자 쿠키가 존재하지 않는다. 그러면 트래킹 사업자 웹서버는 이용자를 식별할 수 있는 값을 새로이 생성하고 (아래 그림 2에서는 "12345" 라는 값이 설정되었다), 이 값을 이용자 웹브라우저에 쿠키를 통해 저장해 둔다. 이용자가 해당 웹서버에 다시 접속하거나 다른 웹사이트에 새로이 접속하면 트래킹 웹서버는 기존에 저장해 놓은 쿠키 값이 존재하는지 조회한다. 만약 쿠키 값이 존재한다면 웹서버는 (1) 자신의 관리하는 데이터베이스에 위 이용자의 인터넷 사이트 방문 이력 정보를 추가로 저장하고 (2) "12345" 번 이용자에 맞는 맞춤형 컨텐츠(광고 등)를 제공한다.

다만, 이러한 트래킹 방식의 중요한 특징은 이용자가 언제든지 웹브라우저의 쿠키를 삭제하거나 리셋(reset)할 수 있다는 점이다. 쿠키가 삭제되거나 새로이 설정되면 그 이용자에 관해 이제까지 트래킹한 정보가 무용지물이 된다. 즉, 그림 2의 세 번째 단계에서 보는 바와 같이 만약 웹브라우저의 쿠키가 삭제되고 나면, 트래킹 웹서버는 마치 해당 이용자가 처음으로 접속한 것과 마찬가지로 인식하게 된다. 따라서 트래킹 웹서버는 새로운 이용자 식별 값(아래 그림 2에서는 "65478")을 지정한다. 그 결과 이전까지 "12345" 이용자에 관해 저장되어 있던 웹사이트 방문 이력은 더 이상 "65478" 이용자로 연결되지 않는다.

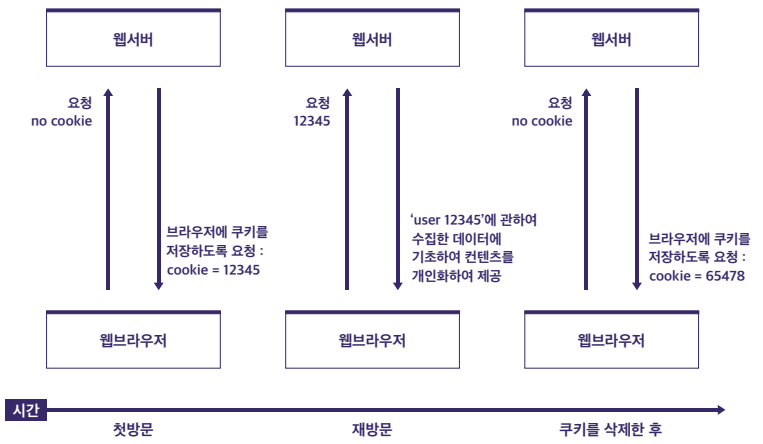

그림 2. 쿠키를 통한 이용자 동일성 식별 과정[5)]

이처럼 이용자의 행태정보를 수집하여 트래킹하는 사업자로는 온라인 맞춤형 광고사업자나 웹사이트 이용통계 조사 업체가 대표적이다. 특히 온라인 맞춤형 광고사업자는 여러 웹사이트에 걸쳐 이용자의 방문 이력이나 검색 쿼리 등의

5) Polonetsky, Jules and Stacey Gray, "Cross Device Tracking: Understanding the State of State Management", Future of Privacy Forum, 2015, p. 3 참조.

정보를 저장함으로써 이용자의 특성을 더욱 정밀하게 파악할 수 있고, 이를 이용하여 이용자의 관심사에 부합하는 맞춤형 광고를 표시할 수 있게 된다. 여기서 주목할 점은 온라인 맞춤형 광고사업자는 웹사이트 제공자와 약정을 통해 허용된 경우에만 해당 웹사이트 방문 이력을 확인할 수 있다는 점이다. 따라서 일반적으로 온라인 맞춤형 광고사업자 입장에서는 가급적 많은 웹사이트 제공자와 약정을 체결하는 것이 유리하게 된다. 한편 많은 웹사이트와 약정을 체결한 광고 사업자에게는 세밀한 분석을 통해 더욱 정밀한 맞춤형 광고를 표시할 가능성이 열리게 된다. 이러한 네트워크 효과의 피드백 구조로 인해 온라인 맞춤형 광고 시장은 소수의 사업자로 집중될 가능성이 크다.

3. 광고 식별자의 개념과 기능: AdID, IDFA 등

광고 식별자란 모바일 운영체제 운영자가 맞춤형 광고 등의 목적으로 별도로 생성한 이용자 아이디를 말한다. 광고 식별자는 모바일 앱에서 주로 활용되는데, 작동 메커니즘은 쿠키의 경우와 유사하다. 가상의 사례를 들어본다. 이용자가 스마트폰에 앱 A, B, C, D를 설치한다고 하자(그림 3 참조). 설치 이후에는 이용자가 각각의 앱을 사용하는 과정에서 서비스 제공자에게 광고 식별자 정보가 보내어진다. 또한 그 과정에서 해당 서비스 제공자가 아닌 광고 사업자 등의 제3자에게 광고 식별자 정보가 보내어지기도 한다. 그러면 이와 같은 제3자는 광고 식별자의 값을 이용하여, 이용자가 어떤 앱을 이용하고 있는지를 포함한 행태정보를 파악할 수 있게 된다. 그림 3의 경우에, 광고 사업자는 광고 식별자 값을 이용하여 동일한 이용자가 앱 A, B, C, D를 이용하고 있다는 사실을 파악하는 것이 가능해 진다. 광고 사업자는 이와 같은 방식으로 여러 앱들 간에 있어서도 이용자를 트래킹할 수 있게 되고, 수집된 정보를 분석하여 맞춤형 광고를 표시하는 것이 가능해진다.

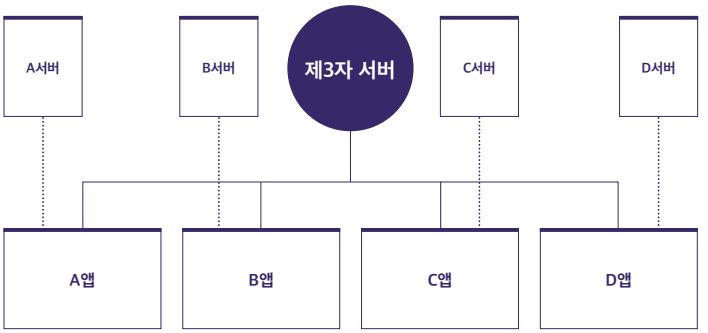

그림 3. AdID 및 IDFA의 작동 방식[6)]

6) 모바일 앱의 소스코드 중 일부 부분에는 AdID 또는 IDFA를 조회하는 내용과 그 밖에 수집된 일정한 종류의 데이터를 연결하여 제3자의 웹서버로 전송하는 내용이 포함되어 있다.

광고 사업자가 이처럼 여러 모바일 앱을 통해 동일한 이용자에 관해 트래킹하는 것이 가능하기 위해서는, 개별 이용자마다 - 좀 더 정확하게는 개별 모바일 기기마다 - 고유하게 부여된 식별자가 존재해야 한다. 그래야 이를 매개로 하여 이용자 정보를 수집하는 것이 가능해진다. 트래킹을 위해서는, 예를 들어, MAC 주소[7], IMEI[8] 등 쉽게 변경될 수 없는 정보나 개인과 쉽게 연결될 수 있는 이메일 주소, 휴대전화번호 등의 정보를 확보하여, 이를 이용자 트래킹용 식별자로 활용할 수도 있다. 다만, 이러한 정보는 변경이 어렵거나 개인과 직접 연결(link)될 가능성이 높으므로 인해 이용자 프라이버시의 침해 위험성이 적지 않다. 더 나아가 이러한 정보를 식별자로 활용하여 이를 매개로 각종 정보를 추가로 수집하게 되면, 이용자의 프라이버시 침해 가능성이 더욱 커지게 된다.

광고 식별자(advertising identifier)는 이러한 이용자 트래킹에 따른 프라이버시 위험을 억제하기 위해 모바일 운영체제 제공자가 개발한 것으로 이해할 수 있다. 이를 구글의 안드로이드 OS에서는 AdID(Advertising ID) 또는 AAID(Android Advertising ID), 그리고 애플의 iOS에서는 IDFA(ID for Advertisers)라 부른다.

AdID와 IDFA는 일종의 표준적인 규칙에 따라 생성되는 총 32개의 알파벳과 숫자, 4개의 하이픈(-)으로 구성된 문자열이다. 개발자들은 모바일 앱의 개발 과정에서 특정한 명령어를 삽입하여 운영체제로부터 이용자가 현재 사용 중인 기기로부터 AdID나 IDFA를 제공 받을 수 있다. 이렇게 확보된 임의적이고 임시적인 AdID나 IDFA를 매개로 하여 이용자에 관한 정보를 수집하게 되는 것이다.

식별자로서의 AdID와 IDFA의 가장 큰 특징은 이용자가 원할 때 언제든 새롭게 변경(리셋)할 수 있다는 것이다. AdID와 IDFA를 매개로 여러 행태정보가 수

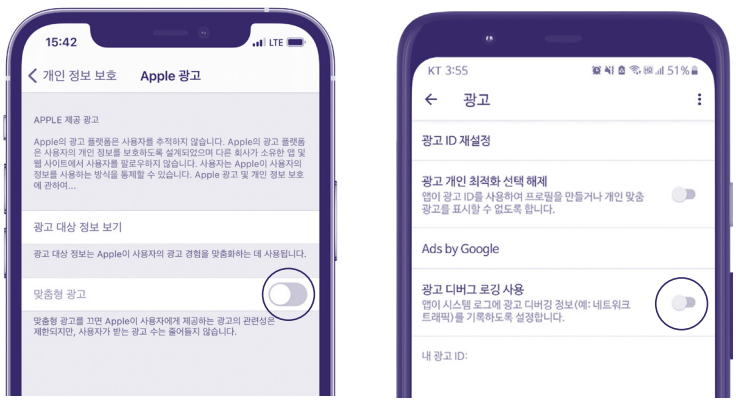

그림 4. 안드로이드 및 iOS의 AdID 및 IDFA 관련 설정 화면[9]

7) Media Access Control Address. IEE 802 네트워크에서 네트워크 인터페이스 컨트롤러(NIC)에 할당되는 고유한 주소이다. 하나의 NIC에는 하나의 고유값이 할당된다.

8) International Mobile Equipment Identity. 3G, LTE, 5G 등의 이동통신 표준에서 휴대전화의 고유 식별자로 활용되는 번호이다.

9) 방송통신위원회·한국인터넷진흥원, 온라인 맞춤형 광고 개인정보보호 가이드라인, 2017, 16면.

집될 수 있지만, 이용자가 이를 변경하면 그 이전까지 수집된 정보와 이용자의 연결은 끊어지게 된다. 이는 웹브라우저 환경에서 이용자가 쿠키를 삭제하거나 새로이 설정하면 그 이전까지 수집된 정보와의 연결이 끊어지는 것과 기능적으로 마찬가지이다. AdID 또는 IDFA와 이용자 사이의 연결 고리를 느슨하게 유지하기 위하여 구글이나 애플은 AdID나 IDFA에 관한 관리 방침[9]을 도입하는 등 일정한 제한을 가하고 있다. 그림 3은 모바일 기기의 설정 메뉴를 통해 광고 식별자의 재설정이 가능한 것을 보여주고 있다.

실제로 웹브라우저 환경에서와 마찬가지로 모바일 앱에서도 여러 개의 앱을 통해 AdID나 IDFA를 수집하는 제3자(트래킹 사업자)가 존재한다. 트래킹 사업자가 모바일 앱 개발 환경에서 보편적으로 사용하는 방식은 SDK(Software Development Kit)를 활용하는 것이다. SDK는 개발자들이 프로그래밍 과정에서 특정한 기능을 구현하고자 할 때 이를 처음부터 직접 개발하는 수고를 하지 않도록 해당 기능을 미리 구현해놓은 코드의 집합을 지칭한다. 개발자들은 특정한 기능을 구현한 SDK를 자신이 개발하는 모바일 앱의 원시코드에 삽입함으로써 그 SDK가 제공하는 기능을 활용할 수 있다.[11]

트래킹 사업자는 모바일 앱을 통해 행태정보를 수집하기 위해 SDK를 주로 활용한다. 예컨대 앱 개발자가 앱 화면에 광고를 표시하기 위해서는 온라인 맞춤형 광고사업자가 제공하는 광고 SDK를 이용한다. 광고 SDK는 해당 앱이 실행될 때마다 함께 실행되어 이용자의 정보를 트래킹 사업자(즉 온라인 맞춤형 광고 사업자)에게 전송한다. 모바일 앱 개발자는 광고 SDK를 통해 광고를 표시하고 광고료를 지급받는다.

이와 같은 방식으로 모바일 앱을 통해서도 이용자에 관한 다양한 행태정보가 수집될 수 있다. 기술적으로는, 웹페이지를 통하는 경우에 비해 모바일 앱을 통해서 이용자 기기에 관해 더 많은 정보를 파악할 수 있다. 모바일 앱이나 SDK는 운영체제를 통해 모바일 기기에 보관된 다양한 정보 또는 각종 센서를 통해 입수할 수 있는 다양한 정보에도 접근할 가능성이 있다. 저장공간에 대한 접근도 웹 브라우저를 이용한 경우에 비하여 상대적으로 자유롭다. 따라서 기술적인 가능성만을 고려한다면, 모바일 앱을 통한 이용자 행태정보 수집은 쿠키를 통한 경우보다 더욱 광범위하게 이루어질 수 있다.[12]

4. 소결

이상에서 본 바와 같이 웹브라우저 환경에서 이용자 트래킹을 위해 제3자 쿠키가, 모바일 앱 환경에서는 AdID나 IDFA와 같은 광고 식별자가 주로 활용되고 있다. 양자의 기본적인 작동 방식은 유사하다. 웹사이트 제작자나 모바일 앱 개발자는 트래킹 사업자에게 이용자 행태정보를 수집할 수 있는 길을 열어 준다. 트래킹 사업자는 이들의 서비스 제공 공간에서 이용자 행태정보를 수집하고, 이렇게 수집한 데이터를 가공, 분석하여 맞춤형 광고 등의 사업을 수행할 수 있게 된

10) 예컨대 AdID나 IDFA를 MAC 주소, IP 주소 등의 다른 식별자와 연결하지 말 것이 대표적이다. "광고 목적: 광고 식별자는 광고 목적으로 영구적인 기기 식별자(예: SSAID, MAC 주소, IMEI 등)에 연결될 수 없습니다. 광고 식별자는 사용자의 명시적 동의가 있어야 개인 식별 정보에 연결될 수 있습니다. 분석 목적: 광고 식별자는 사용자가 명시적으로 동의한 경우에만 개인 식별 정보 또는 영구적인 기기 식별자(예: SSAID, MAC 주소, IMEI 등)에 연결될 수 있습니다." Google, 정책센터(광고), https://support.google.com/googleplay/android-developer/answer/9857753 (2020. 8. 24. 최종방문); 애플도 Apple Developer Program License Agreement를 통해 "use any permanent, device-based identifier, or any data derived therefrom, for purposes of uniquely identifying a device"를 금지하고 있다. Apple, Apple Developer Program License Agreement, 2018. 6., Art. 3.3.9.

11) 예컨대, 압축 파일을 읽어 들여야 하는 모바일 앱을 개발하는 개발자는 시중에 유통되는 압축 파일의 파일 구조를 이해하거나 그 알고리즘을 구현할 필요 없이, 압축 파일에 관련된 SDK를 자신의 원시코드에 삽입하고 해당 SDK에서 정한 명령어를 호출함으로써 압축 파일을 읽을 수 있다.

12) 다른 한편, 쿠키는 간단한 텍스트 파일이어서 개발자나 제3자가 이에 관해 파악하는 것이 상대적으로 수월한 것에 비해, SDK는 개발자 키트 형태로 배포되는 것이어서 개별 앱의 개발자나 제3자가 그 내용을 구체적으로 파악하는 것이 쉽지 않을 수 있다.

다. 트래킹 업체는 다수의 온라인 서비스 제공자를 모집함으로써 많은 양의 행태정보를 축적할 수 있는 기반을 마련한다. 이들은 이러한 정보 수집의 대가로서 광고료를 지급하거나(광고 사업자의 경우), 이용자에 관한 통계 정보를 제공한다(이용 통계 정보 서비스 사업자의 경우). 쿠키든 AdID나 IDFA든 온라인 서비스 제공자가 트래킹 사업자에게 자신의 서비스 공간 일부를 이용할 수 있게 해주고, 트래킹 사업자가 해당 공간에서 이용자에 대한 정보를 수집하도록 허용한다는 점에서 트래킹이 작동하는 방식은 매우 유사하다고 할 수 있다.

III. 국내 데이터 수집 현황

1. 쿠키를 이용한 데이터 수집

국내 인터넷 환경에서의 행태정보 수집에 관한 실증적인 연구는 아직 많지 이루어지지 않았다. 전반적인 국내 인터넷 환경을 대상으로 행태정보 수집에 관해 조사한 연구로는 고학수 외(2013)[13]과 고학수 외(2017)[14]이 있다. 두 연구는 모두 데스크탑 환경과 모바일 환경을 나누어 조사를 하였는데, 아래에서는 데스크탑 환경만을 기준으로 비교한다.

고학수 외(2013)은 국내 주요 웹페이지 61개를 대상으로 쿠키 수집 실태를 조사하였다. 2013년에는 국내에서 인터넷의 이용이 이미 일상화되기는 하였지만, 이 무렵에 쿠키를 통한 행태정보의 수집이 광범위하게 이루어졌다고 평가하기는 어렵다. 조사대상 웹페이지에서 평균적으로 8.4개의 제3자 쿠키가 확인되었다. 쿠키 정보를 수집한 도메인 기준으로는, 가장 많은 웹페이지에서 제3자로서 쿠키를 수집하였던 도메인은 doubleclick.net(Google이 2007년 인수한 온라인 광고 회사) 으로 파악되었다. 이 도메인은 조사 대상 61개 웹페이지 중 절반 이상인 32개 웹페이지에서 총 55개의 제3자 쿠키를 수집하였다. 이어서 10개 이상의 웹페이지에서 제3자로서 쿠키를 수집한 도메인은 criteo.com(19개), nsmartad.com(14개), twitter.com(12개)이 있었다.

4년 뒤에 연구가 수행된 고학수 외(2017)에서는 국내 인터넷 환경에서 이용자 트래킹이 크게 늘어난 것을 확인할 수 있었다. 이 연구는 국내 주요 웹페이지 91개를 대상으로 쿠키 수집 현황을 조사한 것이다. 조사결과 평균적으로 웹페이지당 47.6개의 제3자 쿠키가 발견되었다. 이는 4년 전(2013년)과 비교하여 5배 이상 증가한 값이다. 이 연구에서도 가장 많은 웹페이지에서 제3자로서 쿠키를 수집하였던 도메인은 doubleclick.net 였다. 평균적으로 41.5개의 웹페이지에서 발견되었다. 여전히 주요 웹페이지 중 절반 정도에서 발견된 셈이다. 10개 이상의 웹페이지에서 쿠키를 수집한 제3자 쿠키의 숫자는 4년 전에 비해 크게 증가하여 총 31개에 이르렀다. 그 중에는 4년 전 연구에서 주요한 제3자 쿠키 수집 주체였던 criteo.com(40개)도 있지만, facebook.com(34.8개), widerplanet.com(24개), daumcdn.com(21.7개)과 같이 새로이 상위에 등장한 업체들도 있었다. 특히 페

13) 고학수 외 1명, "국내 인터넷사이트의 개인정보 수집 현황 분석," 법경제학연구, 2013. 12.

14) 고학수 외 2명, 앞의 논문(주 1).

이스북은 2013년 당시에는 1개의 웹페이지에서만 발견되었는데 2017년에는 34.8개의 웹페이지에서 발견되었다. 이를 통해 국내 인터넷 환경에서 페이스북 이용이 증가하면서 이용자 정보 수집 또한 급속히 늘어난 사실을 확인할 수 있었다.

그 외에 국내 인터넷 환경 전체를 대상으로 조사한 연구 결과는 별로 없지만, 특정 산업에 집중하여 행태정보 수집에 관하여 조사한 연구결과는 존재한다. 예컨대, 고은옥(2019)[15]은 국내 12개 온라인 쇼핑몰에서 이루어지고 있는 트래킹에 관해 조사하였다. 이 연구에 따르면, 국내 온라인 쇼핑몰 12개 중 8개에서는 Google Analytics의 트래커가 발견되었고 7개에서는 Criteo, Facebook Custom Audience 등의 트래커가 발견되었다고 한다. 한편 추동균·유진호(2019)[16]는 고은옥(2019)과 동일하게 Ghostery를 이용한 방법론을 활용하여 국내 12개 온라인 쇼핑몰과 8개 신용카드회사 웹페이지에서 발견되는 트래커의 숫자를 조사하였다. 이 연구에 따르면, 신용카드회사 웹페이지에서는 조사 대상이 된 8개 신용카드회사 웹페이지 중 5개에서 Google Tag Manager, 4개에서 Google Analytics, 3개에서 daumcdn.net과 Facebook Custom Audience, WiderPlanet 등의 트래커가 발견되었다. 그리고 온라인 쇼핑몰 12개 웹페이지 중에서는 8개에서는 Google Analytics의 트래커가 7개에서는 Criteo, Facebook Custom Audience 등의 트래커가 발견되었다. 이와 같은 연구는 소수의 한정된 유형의 웹페이지를 대상으로 한 연구이지만, 구글, 크리테오, 페이스북 등이 국내 다수의 웹페이지에서 적극적으로 트래킹을 하고 있다는 점을 확인해 주는 것이다.

한편, 쿠키를 이용하여 수집되는 정보의 구체적인 내용이나 역할이 무엇인지에 관해서는 고학수 외(2017)에서 일부 분석이 이루어졌다. 가장 중요한 정보로는 이용자의 동일성을 확인하기 위한 일련번호가 있다. 이 연구에 따르면 이용자가 91개의 주요 웹페이지를 순차적으로 방문하였을 때, 9번 중 6번은 구글이 저장·조회하는 'IDE 쿠키'의 값이 동일하였는데, IDE 쿠키는 광고 목적으로 활용되는 쿠키로 알려져 있다. 이처럼 쿠기 값이 일정하게 유지되는 동안 구글은 이용자의 동일성을 지속적으로 파악할 수 있다. 페이스북이 광고 용도로 활용하는 'fr 쿠키'의 값도 일반적으로 동일하게 유지되었다. 즉, 이용자가 91개의 웹페이지를 방문하는 동안 브라우저 ID에 해당하는 부분은 1개 또는 2개의 값을 보이며 대체로 일관된 값을 유지하였다. 이처럼 구글과 페이스북은 쿠키를 통해 상당한 수준으로 이용자의 동일성을 확인하고 있는 것으로 조사되었다. 구글과 페이스북은 모두 자사의 웹페이지 등에서 쿠키 등을 통해 웹페이지 방문 내역 등을 수집한다고 공개하였으므로, 이용자의 동일성과 웹페이지의 방문 이력이 함께 수집되는 것으로 이해할 수 있다.

이상의 논의를 요약하자면, 국내 주요 웹페이지를 대상으로 한 연구를 검토한 결과, 소수의 업체가 절반가량의 웹페이지에서 쿠키를 통해 이용자 행태정보를 적극적으로 수집하고 있으며, 수집되는 정보의 내용에는 이용자의 동일성에 관한 정보나 웹페이지 방문 내역 등이 포함되어 있다.

15) 고은옥, 국내 온라인쇼핑몰 트래킹 연결망 분석을 통한 이용자 프라이버시 보호 방안에 관한 연구, 상명대학교 석사학위논문, 2019. 2. 다만, 이 연구는 구체적으로 쿠키를 확인한 연구는 아니고, 트래킹 차단 플러그인인 Ghostery가 차단하였다고 표시하는 트래커의 숫자를 기록한 연구이다.

16) 추동균 외 1명, "국내 신용카드사와 유통회사 웹사이트 온라인 트래킹 비교 분석 연구," 신용카드리뷰, 2019. 12.

2. AdID를 이용한 데이터 수집

AdID와 IDFA를 통한 자료 수집이 어느 정도로 이루어지고 있는지에 대한 국내 연구는 아직까지 찾아보기 쉽지 않다. 해외의 연구 사례는 대체로 IDFA보다는 AdID에 대한 연구에 집중되어 있다. 그 배경에는, 운영체제의 원시코드가 공개되어 있는 구글의 안드로이드와는 달리 원시코드를 원칙적으로 공개하지 않는 애플의 정책상 IDFA를 대상으로 하여 연구를 진행하기에는 기술적인 어려움이 많다는 현실적인 이유가 있다.

최근에 AdID를 통한 국내 인터넷 환경에서의 정보 수집 현황에 대한 연구가 이루어진바 있다.[17] 이 연구의 내용을 간략히 정리하면 다음과 같다. 이 연구에서는 국내의 주요 유료 및 무료 앱 886개를 조사 대상으로 삼아 모바일 앱을 통한 정보수집의 현황에 관해 조사하였다. 이를 위해 조사대상 앱을 휴대전화에 설치하고 구동을 시작하면서 해당 앱의 네트워크 활동을 모니터하는 방식으로 조사를 수행하였다. 이를 통해 기록된 네트워크 활동 자료를 토대로 정규표현식 및 텍스트 검색을 활용하여 전송되는 데이터 내역을 분석하였는데, 특히 AdID, UUID, 이메일 주소, IMEI, MAC 주소, IP 주소 등의 민감성이 있는 정보가 전송되는지에 관해 유의하여 분석이 이루어졌다. 분석의 결과 92.6%에 해당하는 820개 앱이 AdID를 수집하여 서버로 전송하고 있는 것으로 밝혀졌다. AdID는 무료 앱인지 유료 앱인지 여부와 관계없이 수집되고 있었다. 이메일 주소나 IMEI를 수집하는 사례도 일부 파악되었으나 전송 도메인 주소나 전송되는 패킷 내용 등에 비추어 볼 때, 안드로이드 시스템의 운영을 위한 목적으로 전송되는 것으로 추정되었다. 모바일 기기 내부에 저장되어 있는 휴대전화 번호나 주민등록번호 등의 정보가 전송되는 경우는 파악되지 않았지만, 3건의 앱에서 기기 식별성이 있는 MAC 주소를 수집[18]하는 것으로 조사되었다.

한편, 가장 많은 앱으로부터 정보를 수집한 도메인으로 구글, 페이스북, Unity3D(게임엔진업체)를 들 수 있다. 구글과 페이스북 모두 886개의 앱 중 400개를 초과하는 앱에서 AdID를 수집하고 있었다. 이에 비해 Unity3D는 150개 가량의 앱에서 AdID를 수집하였다. 그 이외의 다른 도메인들은 모두 100개 미만의 앱에서만 AdID를 수집하고 있었다. 이처럼 국내 모바일 앱을 통한 AdID 및 관련 정보 수집은 특정 소수의 사업자로 집중되고 있는 현상이 관찰되었다.

AdID와 함께 수집하는 구체적인 정보가 무엇인지에 관하여는 샘플링 조사를 통해 파악하였다. 샘플 조사의 대상으로는 이용자의 연애 상태, 건강 상태, 종교 등을 추단할 수 있는 모바일 앱을 선정하였다. 위 앱들은 이용자의 프라이버시 침해가 현실화할 수 있는 위험이 상대적으로 크다고 판단되었기 때문이다. 위 앱들이 AdID와 함께 전송하는 정보를 조사한 결과, 현재 구동 중인 앱이 무엇인지, 휴대전화 정보(기종 등), 로케일 또는 시간대 정보, 이동통신사 이름, 인터넷 연결상태, 이용자 액션(앱의 설치, SDK의 초기화 등) 등의 정보가 전송되는 것으로 확인되었다. AdID와 이러한 정보들이 지속적으로 결합되어 축적되어 있는 상

17) 김종윤 외 6명, "국내 모바일 앱 이용자 정보 수집 현황 및 법적 쟁점 - ADID를 중심으로," 저스티스(2020. 10. 예정).

18) MAC 주소 수집 사례는 19건 파악되었으나, 그 중 16건은 안드로이드 시스템의 운영을 위하여 수집된 것으로 보인다. 한편, MAC 주소와 AdID가 함께 수집되어 연결될 경우, AdID의 본래적 목적이라고 할 수 있는 '이용자의 직접적인 신원정보와의 단절'이 확보되지 않을 수 있기 때문에 프라이버시 침해 가능성이 커진다.

태에서, 만약 AdID가 이용자의 신원정보와 연결된다면 해당 이용자의 프라이버시 침해 위험이 구체화될 수도 있다.

요약하면, 모바일 앱 환경에서도 구글과 페이스북은 국내 앱의 절반 가량에서 AdID를 수집하고 있는 것으로 나타났고, 그 이외에도 많은 트래킹 업체들이 현재 구동 중인 앱의 이름, 휴대전화 정보, 이동통신사 정보, 이용자 액션 등의 정보를 수집하고 있는 것으로 파악되었다.

IV. 온라인 광고 생태계의 구조

1. 온라인 광고 시장 개관

위에서 본 것과 같이 쿠키나 AdID 등을 이용하여 이용자에 관한 정보를 수집하는 주된 동기 중 하나는 온라인 광고를 제공하는 것이다. 실증 연구를 통해 나타난 바와 같이 쿠키와 AdID 정보를 수집하는 주요 주체는 구글, 페이스북 등의 대형 인터넷 플랫폼 기업이다. 이들 기업들은 매출의 상당부분 - 또는 대부분 - 을 온라인 광고를 통해 발생시킨다. 인터넷 플랫폼을 통한 온라인 광고 시장은 흔히 양면 시장(two-sided market)의 특성을 보인다. 각종 온라인 서비스 제공자(웹사이트, 모바일 앱)들은 많은 경우 이용자들에게 무상으로 서비스를 제공하여 이용자들, 즉 청중들(audience)의 관심을 확보하고자 한다. 그리고 다수 청중의 관심은 광고를 통해 현금화된다. 광고를 통해 얻은 수입은 다시 온라인 서비스 제공자들이 이용자들에게 무상으로 양질의 서비스를 제공할 수 있는 기반이 된다. 그런 점에서 온라인 광고 시장은 인터넷 생태계를 지탱하는 기둥이라고 할 수 있다. 한국온라인광고협회에 따르면, 2019년 국내 온라인 광고 시장은 6조 4213억 원 규모에 이를 정도로 큰 시장으로 성장하였다.[19)]

19) 한국온라인광고협회, 온라인광고 시장 분석 및 전망 2019, 2020. 3., http://onlinead.or.kr/17/?q=YToxOntzOjEyOiJrZXI3b3JkX3R5cXGUiO3M6MzoiYWxsIjt9&bmode=view&idxz=3291169&t=board&category=383QL5Q23o (2020. 8. 21. 최종방문)

2. 온라인 광고의 유형

배너 유형의 온라인 광고 계약은 CPM, CPC, CPA 등의 이니셜로 표현되는 다양한 조건을 통해 이루어진다. CPM(Cost per Mille)은 광고가 천 번 이용자들에게 노출될 때마다 비용을 지급하는 조건, CPC(Cost per Click)은 이용자가 광고를 클릭할 때마다 비용이 지급되는 조건, CPA(Cost per Action)은 이용자가 회원가입, 물품 구매 등 특정한 행위를 할 때마다 비용이 지급되는 조건을 말한다. CPM 조건으로 계약이 이루어질 경우, 광고 사업자는 충분한 이용자만 확보하면 해당 이용자가 어떠한 관심사를 가지는지 신경을 쓸 필요가 없다. 광고의 노출 횟수가 높아지기만 하면 높은 광고비를 받을 수 있기 때문이다. 그런 점에서 이용자의 선호나 특징을 고려한 맞춤형 광고를 할 유인은 높지 않고, 플랫폼 이용자 숫자를 충분히 확보하는 것이 상대적으로 더 중요하다. 이에 비해 CPC, CPA의 조건으로 계약이 이루어지면 광고 사업자는 이용자의 관심사를 정확하게 파악할 유인이 더 커지게 된다. 이용자가 클릭하거나 특정한 행위를 해야 광고비가 지급

되므로, 이용자가 반응을 보일 가능성이 큰 광고를 적절하게 제시하는 것이 중요해지기 때문이다. 이러한 경우 이용자의 관심이나 취향을 정확하게 파악하는 것은 부가가치 창출의 핵심이 된다.

인터넷 플랫폼에서의 광고는 실시간 경매를 이용하는 경우가 적지 않은데, 그런 경우에는 이용자의 관심사에 대한 정확한 파악이 더욱 중요해진다. 실시간 광고 경매 시스템의 개괄적인 작동 방식은 다음과 같다. 우선 이용자가 온라인 서비스 제공자의 서버에 접속 요청을 한다. 그러면 온라인 서비스 제공자는 해당 이용자에 관한 정보를 실시간 광고 경매 시장에 내보낸다. 광고주들은 광고 경매 시장에서 해당 이용자에게 자신의 광고를 보여주기 위하여 지출하고자 하는 비용을 제시하고, 이 중에서 가장 높은 금액을 제시한 광고주의 광고가 이용자에게 전송된다. 이러한 복잡한 절차가 이용자가 웹페이지에 접속하여 웹페이지가 로드되는 1초도 안 되는 짧은 시간 동안 일어난다. 모바일 앱에서도 마찬가지이다. 모바일 앱 이용자가 앱을 작동시키고, 해당 앱의 한 부분에 광고가 게재되는 때까지의 짧은 시간 동안 이용자의 프로파일 정보에 따라 최고의 입찰가를 제시한 광고주의 광고가 낙찰을 받는 과정을 거치게 된다.

실시간 경매 시스템에서 높은 입찰가를 지불하고자 하는 유인은 해당 이용자가 그 광고를 보았을 때 관심을 보일 가능성이 크다는 점에서 비롯된다. 이러한 점에서 특정 광고에 관심을 가질만한 이용자를 최대한 정확하게 식별하는 것은 온라인 광고 생태계에 참여하는 모든 이해관계자들에게 매우 중요한 일이 된다. 이러한 맞춤형 광고에는, 광고주의 웹페이지에 이미 한 번 방문한 이력이 있는 이용자들에게 광고주의 광고를 지속적으로 보여주는 리타게팅(retargetting) 광고와 같이 복잡한 분석기술을 필요로 하지 않는 맞춤형 광고도 있고, 고도의 분석(analytics)을 전제로 하여 이용자의 관심사에 따른 광고를 보여주는 유저 타게팅 광고 등 다양한 유형이 있다.[20)]

3. 온라인 광고 시장의 주요 행위자

온라인 광고 시장에는 매우 다양한 행위자들이 참여하여 복잡한 생태계를 구성하고 있다. 우선 자신의 상품이나 서비스를 광고하고자 하는 광고주가 있고, 다른 한편 광고 공간을 제공하는 인터넷 플랫폼 등의 온라인 서비스 제공자가 있다. 그런데 광고주나 온라인 서비스 제공자 모두 이용자에 관한 충분한 데이터를 가지고 있지 못한 경우가 많다. 그 때문에 온라인 광고 시장은 광고주와 온라인 서비스 제공자 이외에도 매우 다양한 행위자들이 중요한 역할을 수행하게 된다.[21)] 아래 그림 5는 온라인 광고시장의 구조를 개괄적으로 보여주는 것이다.[22)]

그림 5의 가장 왼편에 위치한 것은 광고 공간에 대한 수요자인 '광고주(advertizer)'이다. 그 반대편 가장 오른쪽에 위치한 온라인 서비스 제공자는 광고 공간을 제공하는 공급자로 흔히 '퍼블리셔(publisher)'라고 한다. 온라인 광고 시장은 이들 사이의 수요와 공급을 매칭시키는 역할을 한다. 광고주나 온라인 서비

20) 한국인터넷진흥원, 프로파일링 관련 기술 동향 분석 및 개인정보 정책 방안 연구, 2018, 62~74면.

21) 이하의 구체적인 행위자 유형에 대한 설명은 한국인터넷진흥원, 앞의 자료, 56~61면; Estrada-Jiménez, José et al., "Online Advertising: Analysis of Privacy Threats and Protection Approaches," Computer Communications, Volume 100, 2017. 3. (manuscript version).

22) 이 그림 및 아래에 설명된 역할 구분은 설명의 편의를 위한 것으로 실제로는 하나의 업체가 다양한 역할을 겸하고 있는 경우도 많다.

스 제공자는 이용자의 선호나 성향에 관한 구체적인 프로파일링 정보를 보유하고 있지 못한 경우가 많으므로, 광고주와 퍼블리셔 각각을 위하여 광고에 관한 의사결정을 대신 내려 줄 업체가 필요하게 된다. 이 때 광고주에게 조력을 제공해 주는 업체가 'DSP'(Demand Side Platform)이고, 퍼블리셔에게 조력을 제공해 주는 업체가 'SSP'(Supplier Side Platform)이다. 광고주는 DSP에게 자신이 어떠한 고객을 대상으로 어떤 광고를 하고자 하는지 요구사항을 제공하고, 이를 바탕으로 하여 DSP는 광고주를 위해 광고 시장에 참여하게 된다. 이와 대칭적으로 퍼블리셔는 어떤 특징을 가진 이용자들이 자신의 플랫폼을 이용하는지에 관한 정보를 SSP에게 제공하고, 이에 기초하여 SSP는 퍼블리셔를 위하여 광고 시장에 참여하게 된다.[23)]

23)
이러한 정보의 전달은, 퍼블리셔가 SSP에게 자신의 웹페이지의 공간 일부를 할당하거나, 이들의 SDK를 모바일 앱의 원시코드에 삽입하면서 구체적으로 이루지게 된다.

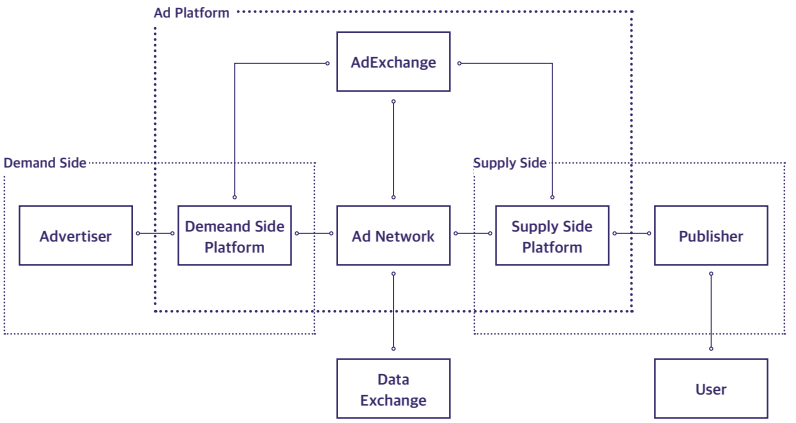

그림 5. 온라인 광고 시장 개요[24)]

24)
Estrada-Jiménez, José et al., 앞의 논문(주 21), p. 4.

DSP나 SSP는 광고의 타겟 매칭이 적절히 이루어지는지 판단하기 위해 'DMP'(Data Management Platform 또는 Data Aggregator라고도 한다)를 통해 이용자들에 관한 정보를 확인한다. 이를 위해 다수의 DMP들로 구성된 'Data Exchange'를 활용할 수도 있다. DSP나 SSP는 DMP 또는 Data Exchange가 보유한 이용자 정보를 활용하여 실시간 광고 공간 경매 시장에 참여하게 되고, 다양한 광고 계약을 체결하기 위한 결정을 내린다. 또한 이들은 이 과정에서 입수하게 된 이용자에 관한 정보를 향후 다른 이용자의 분석을 위하여 사용할 수도 있다.

한편, Ad Exchange는 실시간 광고 경매가 이루어지는 플랫폼을 제공하는 역할도 수행한다. 실시간 광고 경매의 장을 제공하지 않더라도 광고주와 온라인 서비스 제공자들 사이에서 거래의 장을 마련하여 주는 플랫폼은 'Ad Network'라고 부른다.

V.법적 현안 - 결론을 대신하여

앞서 살펴본 것과 같이, 제3자 쿠키 그리고 AdID, IDFA 등을 통한 이용자 행태정보 수집은 국내 인터넷 환경에서 광범위하게 이루어지고 있는 것으로 파악된다. 데스크톱의 웹브라우저 환경과 모바일 앱 환경 모두에서 행태정보의 수집이 일상적으로 이루어지고 있다.

하지만 국내에서는 아직까지 제3자 쿠키나 AdID, IDFA 등에 관한 학술적 연구나 사회적 논의가 충분히 이루어지지 못한 상태이다. 현행 개인정보 보호법상 쿠키, AdID, IDFA 등이 개인정보에 해당하는지 여부는 또렷하지 않다. 개인정보 보호법상 개인정보를 판단하는 주된 기준은 '식별 가능성'이다. 그런데 식별 가능성은 데이터가 수집, 이용되는 맥락에 따라 달리 판단될 수 있어서, 사전적이고 일의적으로 식별 가능성의 유무에 대해 언급하기는 어려운 면이 있다.

국내 선행 연구들에서 조사된 바에 따르면, 트래킹 사업자들이 쿠키, AdID 등을 통해 이용자 행태정보를 광범위하게 수집하고 있는 한편, 개인의 신원에 직접 연결되는 정보를 수집하는 경우는 찾기 어려운 것으로 보인다. 즉, 이들이 수집한 쿠키나 AdID는 여러 데이터와 결합하여 이용자 프로파일링(profiling)에 이용될 수 있지만, 일반적으로 직접 개인의 신원과 연결되는 정보(예컨대, 이메일 주소, 휴대전화 번호 등)와 결합되는 경우는 확인되지 않았다.

다른 한편, 상황에 따라서는 수집된 정보를 이용하여 직접 또는 간접적인 방법으로 개인이 재식별될 수 있는 가능성을 고려해 볼 수 있다. 특히 쿠키나 AdID, IDFA 등은 재설정이 가능한 한편, 쿠키, AdID, IDFA 등의 값이 고정적인 속성값을 가지는 MAC 주소, IMEI 등의 식별자와 결합되거나 개인과의 결합 관계가 비교적 또렷한 이메일 주소, 휴대전화 번호 등의 정보와 결합이 이루어지면 프라이버시 침해 위험이 증가하게 된다. 실제로 구글이나 애플과 같이 AdID, IDFA를 도입한 사업자들은 쿠키나 AdID, IDFA에 식별 가능성이 높은 정보를 연결하여 보관하지 못하도록 막는 정책을 도입하고 있다. 이러한 사업자들의 자발적 정책에 더하여 추가적인 사회적, 제도적 조치가 필요한지 여부에 관하여는 아직까지 많은 논의가 이루어지지 않은 상황이다.

다른 한편, 이러한 프라이버시 침해의 가능성이 있다는 이유만으로 쿠키나 AdID, IDFA의 활용을 터부시할 필요는 없다. 온라인 광고 시장은 인터넷 생태계를 지탱하는 중요한 기둥이고, 이용자 행태정보는 온라인 광고 시장이 높은 부가가치를 창출하는데 있어 중요한 역할을 한다. 또한 행태정보를 활용하여 맞춤형 컨텐츠를 제공함으로써 이용자 효용의 증진이 이루어지기도 한다.

온라인 광고 시장은 이미 자체적으로 프라이버시 침해의 위험성에 대응하기 위한 논의를 진행해오고 있다. 2020년 1월, 구글은 자사 웹브라우저인 크롬(Chrome)으로부터 향후 2년 이내에 제3자 쿠키 기능 지원을 중단하겠다고 밝혔다.[25] 유예 기간 동안 Privacy Sandbox라는 프로젝트를 통해 프라이버시 침해 가능성을 낮추면서 온라인 광고를 할 수 있는 원칙을 모색하겠다며 프로젝트에

25) Google, "Building a more private web: A path towards making third party cookies obsolete", Chromium Blog, 2020. 1. 14., https://blog.chromium.org/2020/01/building-more-private-web-path-towards.html (2020. 8. 24. 최종방문) 참조.

대한 관심을 촉구하였다. 한편, 애플은 최근 2020년 하반기에 적용될 새로운 iOS 14를 공개하면서, IDFA에 관한 정책을 opt-out에서 opt-in으로 변경하겠다고 밝힌바 있다. 앞으로 애플 운영체계에서 구동되는 모바일 앱은 이용자의 동의를 얻어야만 IDFA를 확인할 수 있도록 하겠다는 것이다. 하지만 논란이 커지자 새로운 정책의 이행을 2021년으로 미루겠다고 새로이 정정하여 발표하였다.[26] 이러한 변화와 논란은 온라인 광고 시장의 규모가 커지면서 여러 이해당사자 사이의 입장이 복잡하게 얽혀있는 현실을 반영하는 것이라 할 수 있다. 이러한 시장에서의 변화에 발맞추어, 앞으로 데이터 수집에 관한 연구와 사회적 논의가 보다 활발하게 이루지기를 기대한다.

26) Patience Haggin, "Apple Delays Privacy Change Amid App Publishers' Concerns," The Wall Street Journal, 2020. 9. 3.