

웹로그 분석 시스템의 마케팅 전환 분석 알고리즘에 관한 연구

서민구
 고려대학교 컴퓨터정보통신대학원
 소프트웨어공학과
 e-mail : withsky999@gmail.com

A Study on the Marketing Conversion Analysis Algorithm of Web Log Analytics System

MinGu, Seo
 *Dept. of Software Engineering,
 Korea University Graduate School of Computer & Information Technology

요 약

현재 온라인 마케팅분석을 위해 웹로그 분석시스템을 사용하여 마케팅 결과에 대한 분석을 하고 있지만, 전환에 대한 원인 분석이 사용자의 접속 당시 정보에 의한 데이터로만 분석되고 있어, 본 연구에서는 전환 데이터 분석에 대한 좀 더 현실적으로 분석하기 위한 알고리즘 연구를 제안한다.

1. 서론

최근 웹사이트에서는 웹로그 분석 시스템[1]을 이용한 마케팅 분석을 사용하고 있으며, 특히 온라인 마케팅에 대한 결과를 분석[2]하기 위한 용도로 사용하고 있다.

대부분의 웹로그 분석 시스템에서는 접속자의 유입 경로를 분석해 유입사이트, 검색어 등을 이용해 온라인 마케팅의 결과를 통계적으로 제공해주는 서비스를 하고 있다.

마케팅의 목표인 전환여부에 대한 분석기능은 존재하고 있으나, 단순히 접속건에 대한 전환여부뿐만 아니라 단단하기 때문에, 좀 더 정확한 전환의 원인과 동기를 분석하고자 한다.

2. 사용자 기반 웹로그 분석 시스템

웹사이트에 접속한 사용자를 구분[3]하기 위하여, 쿠키와 세션을 이용하여, 재방문여부를 구분할 수 있다.(동일한 쿠키값인 경우 재방문, 동일한 세션값인 경우 동일한 방문)

본 연구에서는 온라인 마케팅을 하고 있는 웹에이전시 A 사의 2018년 3월부터 2018년 8월까지의 웹로그 데이터를 기준으로 단순 접속에 대한 전환 성공여부 데이터와 새로운 알고리즘 적용 후 전환여부 데이터의 비교하고자 한다.

그림 1은 A 사 웹사이트의 웹로그 데이터 중 일부를 발췌한 데이터베이스 테이블 구조이다.

<표 1> 웹로그 데이터베이스 구조

no	순번
sessionid	쿠키값
connectid	세션값
curdate	접속일시
ip	아이피
refhost	유입경로 도메인
refpage	유입경로 페이지
refparam	유입경로 파라미터
refsearch	유입경로 검색어
motive	요청동기 전환카운트
conversion	전환여부

위 데이터베이스에서 conversion 필드는 전환(견적문의) 성공여부에 대한 값으로 마케팅의 목표가 달성되었는지의 값이다. 이에 대한 시나리오는 사용자가 온라인 마케팅을 통해 웹사이트에 접속 후 여러페이지를 보다가 견적문의 페이지를 통해 개인정보를 남기게 되는 것이다. 즉, 이 A 사의 웹사이트에서 온라인 마케팅을 통해 웹사이트에서 사용자를 전환(예:견적문의)을 일으키기 위한 수단이 된다.

3. 요청동기 접속

현재의 웹로그 분석 시스템에서는 웹사이트를 접속한 사용자의 유입경로 정보에 따라 전환여부가 결정되는 형태로, 일자별, 유입경로, 검색어 별로 전환여부에 대한 통계를 추출해 낼 수 있다.

하지만, 위의 경우와 다른 시나리오에 대해서는 시스템에서 체크하지 못한다. 예를 들어, 1월 1일 키워드 마케팅을 하고 있는 네이버에서 ‘웹사이트제작’이라는 검색어로 접속하여 해당 웹사이트를 즐겨찾기하고, 다음날 1월 2일 북마크를 통해 해당 웹사이트를 방문하여 전환이 일어난 경우 이 접속은 마케팅을 통한 전환이 아닌, 직접방문한 사용자의 전환으로 인식되는 경우가 발생한다. 이 전환 건은 마케팅의 결과로 전화이 발생한 건이지만, 실제 웹로그 분석에서는 직접방문으로 발생한 전환으로 인식하게 된다.

따라서, 위의 예처럼 마케팅을 통한 접속으로 해당 웹사이트를 인지한 후 직접 재접속 하는 경우도 마케팅의 결과로 인한 전환으로 데이터를 처리하는 알고리즘이 필요하다.

4. 기존 데이터 분석

아래 표 2는 A사의 2018년 3월 1일부터 2018년 8월 31일까지의 웹로그 데이터 26,453건 중 전환이 이루어진 데이터에 대한 수치이다.

<표 2> 전환데이터 유입구분 별 전환 수

사이트	전환수
직접방문	95
네이버	115
다음	26
합계	236

총 접속 데이터 26,453건 중 전환이 이루어진 유입건수는 236건이며, 이중 40%인 95건이 직접 방문하여 전환이 이루어진 것으로 나타났다. 여기서 직접방문이란 웹사이트 방문 시 레퍼러(이전 접속 정보) 값이 없거나, 온라인 마케팅을 하고 있는 검색엔진(예: 네이버, 다음, 구글 등)을 통한 접속이 아닌 건을 말한다.

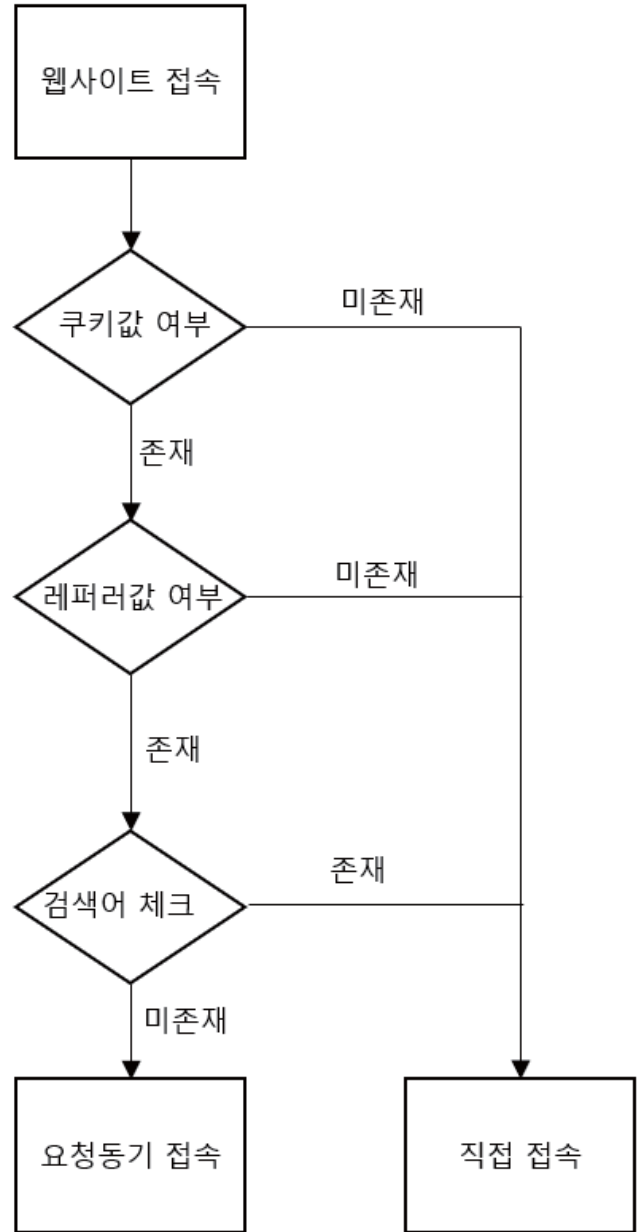
5. 요청동기 분석 알고리즘

전환 당시는 직접방문으로 접속하였지만, 예전에 마케팅으로 접속했던 내역이 있다면 이 방문은 마케팅 접속으로 인한 전환으로 분석하고자 한다.

알고리즘 순서는

1. 현재 접속의 쿠키값과 동일한 접속 내역이 존재하여야 하고,
2. 이전 접속의 레퍼러값이 존재하여야 하며,
3. 이전 접속 시 검색한 단어가 A사의 웹사이트를 알고 검색한 것이라고 유추할 수 있는 단어가 아니어야 한다.

위 알고리즘 1,2,3 모두 만족하는 경우 요청동기를 갖고 있는 접속으로 판단해 마케팅으로 인한 접속이라고 분류한다.



(그림 1) 요청동기 분석 알고리즘 순서도

그림 1은 요청동기 분석 알고리즘을 순서도로 나타낸 것이며, 알고리즘 중 3번은 이전 접속 시 검색한 단어가 A사를 이미 알고 검색한 단어인지 구별하는 방법이다. 예를 들어, 삼성 웹사이트를 들어가기 위해 ‘삼성’이라는 검색 키워드로 검색 후 접속하는 경우는 마케팅의 효과라고 보기 어렵기 때문이다.

따라서 사전에 데이터베이스에 테이블을 만들어 요청동기 제외 검색어를 저장해 이전 접속에 해당 단어를 검색어로 유입된 경우는 요청동기 유입에서 제외했다.

```
SELECT sessionid, connectid, refhost, curdate, refsearch
FROM url
WHERE sessionid='1526013452848'
```

sessionid	connectid	refhost	curdate	refsearch
1526013452848	1526013452848	search.naver.com	2018-05-11 13:37:25	
1526013452848	1526264036959	search.daum.net	2018-05-14 11:13:47	

(그림 2) 요청동기 접속으로 인한 전환 데이터

그림 2는 요청동기 접속의 실제 데이터로 2018년 5월 11일 13시 37분에 네이버를 통해 최초 접속했으며, 이후 3일 후인 2018년 5월 14일 11시 13분에 다음을 통해 재접속한 것을 알 수 있다.

동일한 sessionid (쿠키값), refhost (레퍼러 호스트)가 존재하며, 검색어는 회사명이 아닌 단어로 접속해 전환이 일어난 경우이다.

6. 요청동기 분석 알고리즘 적용 후 데이터 분석

앞서 직접방문은 95건의 전환이 발생했지만, 요청동기 분석 알고리즘을 적용한 후 동일한 데이터를 분석해, 아래 표 3과 같은 결과를 얻었다.

<표 3> 요청동기 접속수별 전환 수

요청동기 접속수	전환수
0	3
1	87
2	2
3	1
4	1
7	1
합계	95

표 3의 요청동기 접속 수는 마케팅 접속에 의한 누적 접속 수이며, 요청동기 접속 수가 0인 전환 수 3건을 제외한 92건은 실제 직접방문이 아닌 마케팅 접속의 결과로 인한 전환임을 알 수 있다.

7. 결론

온라인에서 웹사이트를 운영하며, 마케팅을 하고 있는 대부분의 업체는 상당히 많은 마케팅비용을 지

불하며, 최소한의 비용으로 최대한의 효과를 얻기 위해 노력하고 있다. 마케팅 효과를 개선하기 위한 수단으로 웹로그 분석 서비스를 이용하여, 광고매체나 검색키워드 등을 조절하고 있다.

기존 웹로그 분석 시스템에서는 접속 데이터에 대한 분류, 마케팅 유입 경로, 전환여부 등에 대한 표면적인 데이터로 분석하였지만 위 결과에서 보듯, 실제 수치상의 통계보다 더 높은 마케팅 효과가 있음을 알 수 있었다.

이 연구가 마케팅의 결과인 전환을 좀 더 정확히 분석하기 위한 방법의 하나로 사용되기를 바라며, 향후 연구에서는 더 나아가 사용자 접속패턴이나 페이지 이동경로 등을 분석해 전환이 일어나지 않은 건들의 원인 분석에 도움이 될 것으로 사료된다.

참고문헌

- [1] 아비나쉬 카우쉬. “실전 웹사이트 분석 A to Z : 성공적인 온라인 마케팅을 위한 웹데이터 분석과 활용”
- [2] 김영철. “웹로그 데이터 분석을 위한 마케팅 패러다임 연구”
- [3] 장경호. “효율적인 웹 마이닝을 위한 정확한 사용자의 행동 패턴 포착”