

키워드 광고 시스템의 설계 및 구현

지혜성*, 류기곤*, 임희석*
*한신대학교 소프트웨어학과
e-mail: starsmania@hanmail.net

Design and Implementation of Keyword Search Advertising System

Hyesung Jee*, Kigon Lyu*, Heuseok Lim*
*Hanshin University. Dept. of Software.

요 약

본 논문은 포털 사이트의 효율적인 광고 제공을 위한 자연어처리 기반의 키워드 광고 시스템을 제안한다. 사용자의 질의에 대한 형태소 분석 결과를 사용하여, 기존의 키워드 정합에 의한 광고 시스템보다 재현율을 향상시킬 수 있었다. 또한, 웹 기반 키워드 광고 뿐 아니라 메시지를 통한 대화 내용 기반 키워드 광고도 제안한다.

1. 서론

포털 사이트는 인터넷을 자유롭게 돌아다닐 수 있도록 도와주는 역할을 담당한다. 하루에도 많은 사람들이 포털 사이트에 접속하여 정보를 검색할 뿐 아니라 다른 사람들과 콘텐츠나 정보를 공유한다. 이처럼 사람들의 활용도가 높은 포털 사이트는 광고에 대한 기대효과가 좋은 환경을 지니고 있어, 많은 광고주들의 홍보 수단이 되고 있다.

포털 사이트의 광고는 더 이상 낫설지 않은 모습이 되었으며, 현재는 수많은 광고들이 삽입되어 있어, 사용자들의 집중도가 떨어지고 시선이 분산되는 단점이 나타나고 있다. 포털 사이트를 이용한 광고가 서로 경쟁하면서, 더 효율적인 광고를 고안하기 위해 동영상을 이용하거나 사용자의 참여를 유도하는 광고가 등장하였다. 그러나 이러한 광고들은 사용자의 관심사와 상관없이 일방적으로 제공되기 때문에, 사용자로부터 외면당하는 경우가 적지 않다.

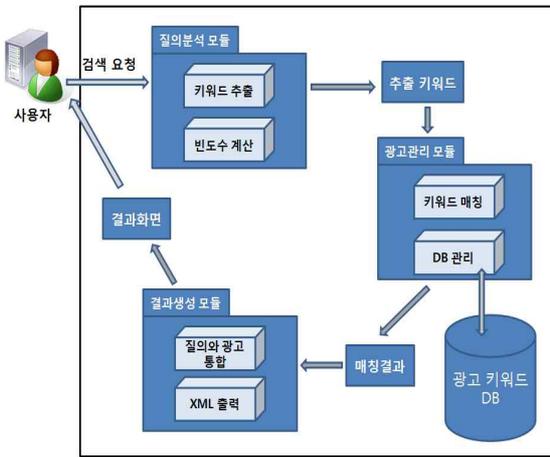
최근에는 사용자의 관심사를 고려한 키워드 기반의 광고가 제공되고 있다. 이는 포털 시스템을 통해 사용자가 정보를 검색할 때, 사용자의 질의를 분석하여 질의 내 키워드와 연관된 광고를 제공하는 방식이다. 키워드 광고는 사용자의 정보 요구가 질의

로 표현된다는 점을 이용하여, 질의와 연관된 검색 결과 뿐 아니라 기업이나 상품에 대한 광고를 제공하는 것이다. 키워드 광고는 현재 포털 사이트에서 가장 많이 사용되고 있는 광고 방식 중 하나로서, 동일 키워드에 대한 광고 순위가 중요하다.

본 논문에서는 자연어처리 기반의 키워드 광고 시스템을 개발하여, 단순 키워드 정합에 의한 기존 키워드 광고 시스템의 재현율을 향상시키고, 웹 기반 키워드 광고뿐 아니라 대화내용 기반 키워드 광고도 제안한다.

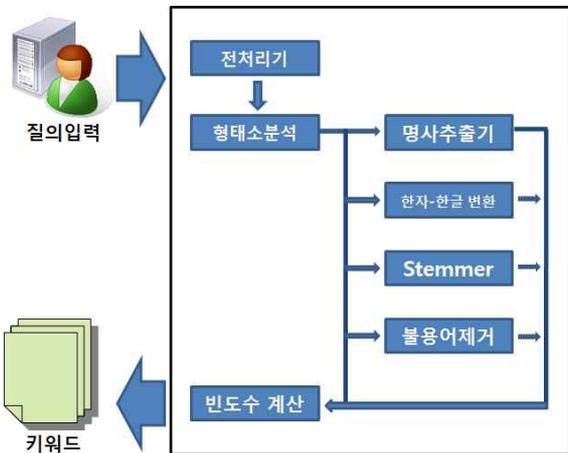
2. 키워드 광고 시스템

본 논문에서 제안하는 키워드 광고 시스템은 [그림 1]과 같다. 질의를 어절 단위로 형태소 분석하여 키워드를 찾아내는 질의분석모듈과 광고주의 광고를 관리하고 키워드와 연관된 광고를 검색하는 광고모듈, 질의와 연관된 광고를 검색 결과에 추가하는 결과생성모듈로 구성된다.



[그림 1] 키워드 광고 시스템 구조도

2.1 질의분석 모듈



[그림 2] 질의분석 모듈

질의분석 모듈은 [그림 2]와 같이 형태소분석을 통해 사용자로부터 입력받은 질의 내 키워드를 추출한다.

형태소분석 과정은 전처리, 불용어 제거, 그리고 명사추출로 이루어진다. 전처리과정은 사용자 질의에 포함되어 있는 한글, 영어, 한자, 숫자, 특수문자 및 문장부호 등을 각각 분리한다. 그리고 한자는 한글로 변환하고 영어는 어간을 추출하여 키워드 후보로 저장한다. 불용어 제거 과정은 전처리를 통해 분리된 한글 어절 중 기 구축된 불용어 사전을 통해 불용어로 판명된 어절을 제거한다. 마지막으로 명사추출은 기 구축된 명사추출기를 통해 불용어가 제거된 한글 어절에서 명사를 추출하고, 이를 키워드 후보로 저장한다. 본 논문에서 명사를 키워드 후보로 추출한 이유는 대부분의 광고 키워드가 명사이기 때

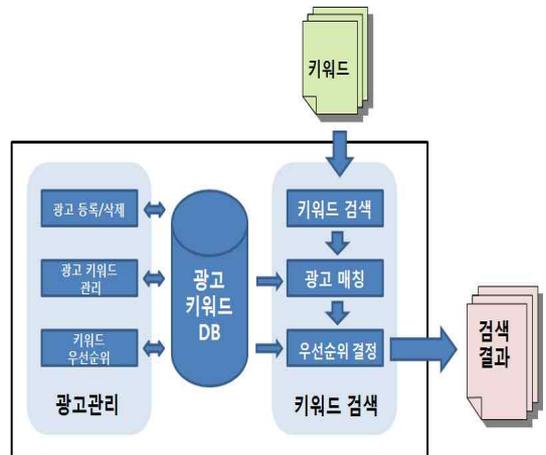
문이다.

질의의 특성 상 포함되어 있는 단어의 길이가 길지 않고 수가 많지 않음을 고려하여, 질의분석을 통해 추출된 키워드 후보의 빈도수를 각각 계산함으로써 질의 내 키워드의 가중치를 계산할 수 있다. 이는 질의 내 여러 키워드 후보 중 적합한 키워드 광고를 찾기 위해 사용된다. [그림 3]은 질의분석을 통한 키워드 가중치 결과를 보여준다.

질의: 한신대학교 어떻게 가나요?
 전처리: 한신대학교, 어떻게, 가나요
 명사추출: 한신대, 한신대학교, 학교
 불용어제거: -
 빈도수 계산: 한신(3), 한신대(2), 한신대학교(1), 학교(1)

[그림 3] 질의분석을 통한 키워드 가중치 결과

2.2 광고 모듈



[그림 4] 광고모듈

광고 모듈은 [그림 4]와 같이 광고주의 광고를 관리하는 광고관리와 질의 내 추출된 키워드와 연관된 광고를 검색하는 키워드검색으로 구성된다.

광고 관리는 광고주가 구매한 키워드를 관리하고 동일 키워드에 대한 우선순위를 조절한다. 키워드검색은 질의분석 모듈에서 추출된 키워드로 광고DB를 검색하여 키워드와 연관된 광고 내용을 검색결과로 제공한다. 광고에는 3개까지의 키워드를 등록할 수 있으며, 광고는 텍스트와 이미지, 동영상(Flash)으로 구분된다. 이러한 정보를 넣을 DB 테이블의 구성은 <표 1>과 같이 설계 하였다.

advertising_task_list	
PK	task_id
	advertiser_id
	item_id
	task_type
	assigned_keyword
	start_date
	expiration_date
	request_count
	is_expiration
	is_delete
	registration_date
	modify_date

advertising_items	
PK	item_id
	advertiser_id
	item_type
	target_url
	text_title
	text_detail
	image_src
	image_width
	image_height
	is_delete
	registration_date
	modify_date

keyword_type	
PK	keyword_item_id
	task_id
	advertiser_id
	keyword_id_1st
	keyword_id_2nd
	keyword_id_3rd
	is_delete
	registration_date
	modify_date

keyword_repository	
PK	keyword_id
	keyword
	is_delete
	registration_date

<표 1> DB 테이블 구성도

2.3 결과생성 모듈

결과생성 모듈은 질의 내 추출한 키워드와 관련된 광고를 XML 형태의 정보로 가공하여, 정보검색 시스템의 결과에 삽입하도록 도와준다.

검색된 광고는 검색 결과의 최상단에 위치한다. 생성된 광고의 XML 정보는 advertiserID(광고주명),

type(광고 종류), targetUrl(광고와 관련된 주소), textTitle(광고 제목), textDetail(광고 내용) 및 이미지나 동영상 광고의 속성 값(파일 이름, 가로 사이즈, 세로 사이즈)으로 구성된다.

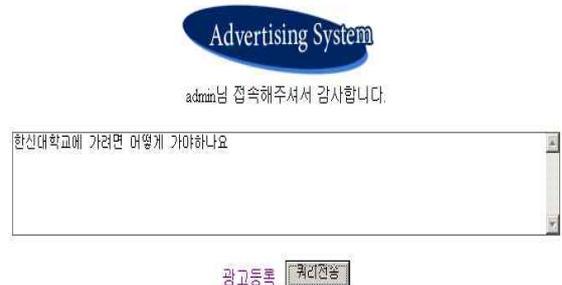
[그림 5]는 추출한 키워드와 관련된 광고를 XML 형태로 생성한 결과이다.

```
<?xml version="1.0" encoding="UTF-8" ?>
<taskList>
  <title>keyword Advertising</title>
  <item>
    <title>5</title>
    <advertiserID>admin</advertiserID>
    <type>img</type>
    <targetUrl>http://www.hs.ac.kr</targetUrl>
    <textTitle />
    <textDetail />
    <imageSrc>hanshin.gif</imageSrc>
    <imageWidth>0</imageWidth>
    <imageHeight>0</imageHeight>
    <swfSrc>untitled.gif</swfSrc>
    <swfWidth>0</swfWidth>
    <swfHeight>0</swfHeight>
    <freq>0</freq>
  </item>
```

[그림 5] XML 형태로 생성된 광고 정보

3. 실험 및 결과

제안하는 시스템의 성능 평가를 위해 정확율과 재현율을 측정해야 하지만, 포털 사이트와 동일한 데이터에 대한 비교 실험이 불가능하기 때문에 정량적인 실험 결과를 도출하기 어렵다. 따라서 본 논문에서는 정확한 성능 평가 대신 시뮬레이션 결과를 실험 결과로 대치하고, 정량적인 성능 평가는 향후 연구 과제로 삼는다.



[그림 6] 키워드 광고 시스템 입력 화면

제안하는 시스템은 [그림 6]과 같은 검색 화면에 서질의뢰를 입력받아 [그림 7]처럼 검색 결과에 키워드 광고를 삽입하였다.



[그림 7] 검색 결과 화면

본 논문은 웹기반 키워드 광고 뿐 아니라 [그림8]처럼 대화내용에 기반 한 키워드 광고를 제안한다. 메신저의 대화창이나 쪽지 기능을 통해 대화를 주고, 받을 때 대화내용을 질의로 분석하여 대화내용과 관련된 광고를 제공한다.

이는 사용자의 기호를 고려한 사용자 맞춤형 광고를 제공함으로써 높은 광고 효과를 기대할 수 있다.



[그림 8] 대화내용기반 키워드 광고

4. 결론 및 향후 연구 과제

그동안 키워드 광고는 다른 인터넷 광고보다 마케팅 광고 효과 등 여러 가지 측면에서 높은 기대효과를 인정받으며 지속적인 성장을 해왔다. 그러나 사용자의 기호나 관심사를 고려한 키워드 광고에 대한 연구가 많지 않았다.

이에 본 논문은 효율적인 키워드 광고를 위해 자연어처리 기반의 키워드 광고 시스템을 설계하였고, 대화내용 기반의 사용자 맞춤형 광고도 제안하였다.

향후 정량적인 성능 평가를 통하여 효율적인 키워드 광고 시스템으로 개선하고, 웹기반 광고 뿐 아니라 대화내용 기반의 사용자 맞춤형 키워드 광고를 연구하여 효과적인 광고 시스템을 연구할 것이다.

참 고 문 헌

- [1] 김명철, 김덕봉, 김유성, 김재훈, 박혁로, 이하규 공역, 최신 정보 검색론 2001.
- [2] 김호, MySQL로 배우는 데이터베이스프로그래밍 2006.
- [3] JAVAJIKI.NET, Model1, Model2, Struts 개발 방식의 비교 분석.
- [4] 홍성용, XML원리와 응용 2003.
- [5] 최학원, “전문화된 키워드 검색광고 서버와 관련 시스템 설계 및 구현”, 성균관대학교 정보통신대학원, 2007.
- [6] 김도연, “키워드 검색 광고에서의 산업별 CPC와 CPM 방식에 따른 광고 효과에 관한 연구”, 한양대학교 경영대학원, 2008.
- [7] 주정인, “Local Search Advertisement Business model”, 高鳳論集, Vol.39 No.-[2006], 33-48, 2006.