
사용자 질의 의미 해석을 위한 선호도 기반 검색 기술

정훈*, 이무훈**, 도하니***, 최의인****

Preference-based search technology for the user query semantic interpretation

Hoon Jeong*, Moo-Hun Lee**, Hana Do***, Eui-In Choi****

요약 질의 의미 해석에 대한 대표적인 시맨틱 검색은 논리적으로 표현된 지식 베이스를 사용하여 현재의 키워드 기반 검색보다 더 정확한 결과를 제공할 수 있다. 기존의 키워드 기반 검색 시스템은 사용자 키워드의 의미가 상호 연결이 되어 있지 않아서 사용자의 질의 의미 해석을 위한 선호도 검색을 할 수가 없다. 본 논문에서는 사용자의 검색 의도에 부합하는 정확한 검색 결과를 제공할 수 있는 사용자 검색 선호도 기반으로 평가하여 랭킹하는 기법을 제안한다. 제안한 기법은 온톨로지 기반으로 구축된 지식 베이스의 정형화된 구조에 의미 해석 과정이 통합된 온톨로지 지식 베이스 기반 검색 시스템이다.

주제어 : 시맨틱 검색, 온톨로지, 사용자 선호도, 쿼리

Abstract Typical semantic search query for Semantic search promises to provide more accurate result than present-day keyword matching-based search by using the knowledge base represented logically. Existing keyword-based retrieval system is Preference for the semantic interpretation of a user's query is not the meaning of the user keywords of interconnect, you can not search. In this paper, we propose a method that can provide accurate results to meet the user's search intent to user preference based evaluation by ranking search. The proposed scheme is Integrated ontology-based knowledge base built on the formal structure of the semantic interpretation process based on ontology knowledge base system.

Key Words : Semantic Search, Ontology, User Preference, Query

1. 서론

정보검색(Information retrieval)은 자료를 수집/분석/가공하여 축적하고, 축적된 자료로부터 사용자가 원하는 데이터/정보/지식 등을 찾아주는 일련의 과정으로 정의한다. 정보 검색은 인터넷과 같은 기반 IT 기술이 발전하면서 특정 도메인의 전문가 집단에서 사용되는 전문물이 아닌 일반 사용자들이 쉽게 접근하고 이용할 수 있는 대표적인 인터넷 서비스로 자리 잡고있다.

최근에 네이트나 네이버에서 새롭게 시도하고 있는

시맨틱 검색은 사용자의 의도와 검색어의 내포된 뜻을 기반으로 여러 종류의 검색 결과를 제시하는 방식으로 지능형 검색 또는 의미기반형 검색이라고 말할 수 있다.

시맨틱 검색은 시맨틱 웹의 개념에서 파생된 결과물이라고 할 수 있다[1,2].

본 논문에서는 사용자 검색 키워드를 지식 베이스 기반으로 의미해석 하였다. 그리고 사용자 검색 의도에 부합하는 정확한 검색 결과를 제공할 수 있는 사용자 검색 선호도 기반으로 평가하여 랭킹하는 기법을 제안한다. 이렇게 랭킹된 의미해석 결과를 검색 키워드 형태로 사

*한남대학교 컴퓨터공학과 석사과정

**한국전자통신연구원

***한남대학교 컴퓨터공학과 석사과정

****한남대학교 컴퓨터공학과 교수(교신저자)

논문접수: 2013년 1월 17일, 1차 수정을 거쳐, 심사완료: 2013년 2월 13일, 확정일: 2013년 2월 20일

용자에게 추천해주고 사용자에게 선택할 수 있게 함으로써 좀 더 정확한 검색 결과를 제공할 수 있도록 하였다.

2장에서는 정황 정보 기반 검색 및 질의 의도 해석 기반 검색에 대한 관련연구를 소개하였고, 3장에서는 질의 그래프 기반 키워드 해석, 4장에서는 사용자 키워드 의미 해석 및 평가모델, 5장에서는 사용자 선호 기반 평가 모델, 6장에서는 질의 의미 해석을 위한 사용자 선호 기반 평가 모델의 성능 평가를 설명하였다. 마지막으로 7장에서 결론을 맺는다.

2. 관련 연구

2.1 정황 정보 기반 검색

상황 정보란 유비쿼터스 컴퓨팅과 관련하여 사용자와 다른 사용자, 시스템, 혹은 디바이스의 애플리케이션 간 상호 작용에 영향을 미치는 사람, 장소, 사물개체, 시간 등 상황의 특징을 규정하는 정보를 의미한다. 상황인식 컴퓨팅 기술은 사용자의 현재 위치, 시간, 주변에 있는 다른 사람이나 정보가진 기기들, 사용자의 행동 및 작업 이력 등과 같은 사용자의 현재 상황정보를 파악하고 분석하여 사용자가 현 상황에서 필요로 하는 서비스를 검색하여 구동시켜 주는 기술이다. 상황인식 컴퓨팅 환경에서 애플리케이션들은 상황인식 모델을 기반으로 개발되고 실행되는데 상황인식 시스템에서는 상황정보를 가공하고 저장하여 공유하기 위한 효율적이며 통합된 상황인식 모델을 제공할 수 있어야 하여 이를 위하여 상황인식 시스템은 적절한 상황인식 모델을 구성하고 이를 관리할 수 있는 기능과 능동적이고 지능적인 서비스를 위하여 상황인식 모델에 명시적으로 표현되지 않은 암시적인 상황정보를 추론할 수 있는 기능을 제공해야 한다[3]. 상황 정보 추론은 다양한 센싱 데이터를 융합하여 상위 상황 정보를 유도하기 위해 확률적인 메커니즘을 제공하여 추론을 하는 것이 중요하다. 온톨로지 기반 추론 엔진은 온톨로지로부터 얻은 지식과 사실을 유추하는 기능을 실행한다. 즉 온톨로지를 대상으로 추론 엔진을 통해서 질의를 돌려주게 된다. 이때 추론엔진은 질의어의 관계를 파악하여 관련 용어를 검색 엔진에 전달하게 된다. 이러한 상황정보 추론(context information reasoning)에 대한 대표적인 연구로는 Hoolet, F-OWL, Jena 등이 있다.

2.2 질의 의도 해석 기반 검색

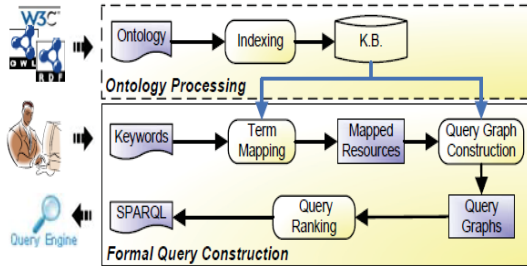
웹 질의의 숨겨진 추상적인 사용자 의도를 파악하려고 한 대표적 연구로 웹 질의의 사용자 의도를 Navigational, Informational, Transactional 질의로 나누는 것을 제안하였으며 이 연구는 사용자 검색 목적에 관련된 다양한 연구를 촉발시켰고[4] 이 중에서 Broder의 분류체계(taxonomy)를 더욱 세분화하고 3개의 상위 범주와 8개의 하위 범주를 통한 분류체계를 제안하였다[5]. 그러나 이전 연구들의 한계는 질의의 의도가 매우 추상적으로 표현되어 실제 검색환경에서 질의들에 나타나는 사용자의 다양한 의도들을 충분히 표현하지 못했다. 한편 사용자의 질의 의도를 규칙기반과 통계 기반의 방법을 접목시킨 하이브리드 방법을 이용하여 105개의 시맨틱 카테고리에 분류하는 연구와[6] 질의의 사용자의 의도를 문서분류를 통한 개념(class)으로부터 언어내려 하는 연구가 있었다[7].

3. 질의 그래프 기반 키워드 해석

SPARK는 키워드 기반 검색에 익숙한 사용자들이 입력하는 키워드를 질의 그래프 기반으로 의미를 해석하고 해석된 결과로부터 시맨틱 검색을 위한 정형화된 질의어를 구성하여 검색 결과를 제공하는 시맨틱 검색 프레임워크이다[8]. 그림 [1]은 SPARK 프레임워크의 구조를 설명하고 있다. SPARK는 개체 식별, 질의 그래프 생성, 질의 랭킹의 세 단계를 통해 사용자 검색 키워드의 의미를 해석하고, SPARQL 형태의 정형화된 질의를 생성하여 검색 결과를 제공한다. SPARK는 키워드 기반 검색에 익숙한 사용자들이 정형화된 질의의 표현 방법을 전혀 모르는 상태에서 일반 키워드 검색과 동일한 검색 키워드를 제공하고 SPARK는 이를 해석하여 최적의 질의 조합을 구성하는 것이 특징이다.

- 개체 식별(Term Mapping): 사용자 키워드로부터 지식 베이스 내의 지식 개체를 식별
- 질의 그래프 생성(Query Graph Construction): 식별된 지식 개체를 바탕으로 시맨틱 네트워크를 탐색하여 질의 그래프를 구성
- 질의 랭킹(Query Ranking): 생성된 질의 그래프를 사용자 질의와 지식 베이스를 기반으로 평가하여

사용자 검색 의도에 가장 부합하는 순서로 질의 그래프를 순위화

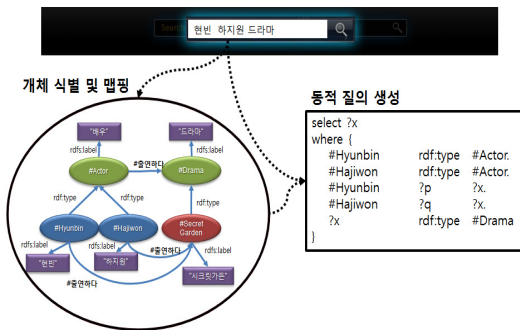


[그림 1] SPARK 프레임워크

4. 사용자 키워드 의미해석 및 평가모델

4.1 질의 의미해석(Query Interpretation)

식별하는 과정에서 개체 식별이 완료되면, 지식 베이스로부터 식별된 개체를 모두 포함하는 부분 그래프를 찾는 질의 의미해석을 수행한다. 부분 그래프를 찾는 과정에서 식별된 지식 개체를 포함하는 부분 그래프는 다양하게 존재 가능하다. 이러한 부분 그래프는 사용자가 입력한 검색 키워드로부터 해석 가능한 다양한 의미해석 후보가 될 수 있으며, 이러한 의미해석 후보를 지식 베이스 기반으로 평가하여 가장 적합한 질의 의미해석을 결정하고, 결정된 의미해석으로부터 SPARQL 형태의 질의문을 생성한다.



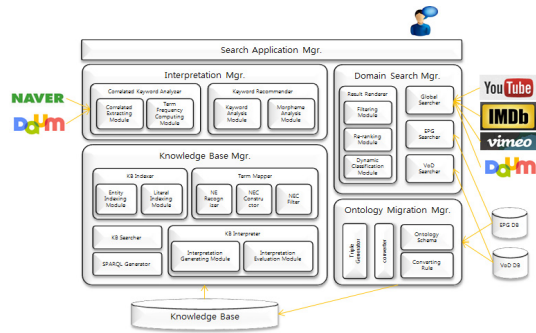
[그림 2] 질의 의미해석 절차

[그림 2]에서 보는 바와 같이 “현빈 하지원 드라마”에 대한 지식 개체 식별이 완료되면, #Hyunbin, #Hajiwon으

로부터 #Actor 클래스를 찾고 #Actor 클래스로부터 #Drama로 연결되는 부분 그래프를 탐색한다. 그림에서 설명하는 온톨로지는 단순한 형태의 온톨로지로서 표현되었지만, 일반적으로 구축된 온톨로지에서는 풍부한 메타 데이터를 바탕으로 다양한 형태의 연결 관계를 뽑아 낼 수 있게 된다. 심지어 인스턴스 간의 관계 탐색 깊이가 깊어질수록 무수히 많은 부분 그래프를 뽑아낼 수 있게 된다. 논문에서는 개체 식별 결과로부터 가능한 모든 부분 그래프를 뽑는 것에 초점을 두었다. 논문의 다음 절에서 설명하는 질의 의미해석 평가 모델에 따라 생성된 의미해석 후보들을 평가하여 랭킹하면 사용자 검색 의도에 만족하는 검색 결과를 제공할 수 있기 때문이다.

4.2 질의 의미해석 평가 모델

[그림 3]은 온톨로지 지식 베이스 프레임워크 구조이다.



[그림 3] 온톨로지 지식 베이스 프레임워크 구조

이 프레임워크 구조를 보면 사용자가 입력한 검색 키워드에 대한 의미해석 결과는 개체 식별 단계에서 1:n개의 개체를 식별하고 식별된 개체들의 조합을 생성하여 지식 베이스 상의 부분 그래프를 탐색하게 되면 다시 n:m 개의 의미해석 결과가 생성된다. 특히 지식 베이스가 얼마나 많은 양의 정보를 포함하고 있는냐에 따라 무수히 많은 의미해석 결과를 생성할 수 있다. 따라서, 논문에서는 사용자 검색문에 대해 가능한 모든 의미해석 결과를 생성하고, 생성된 의미해석 결과를 제한하는 평가 모델을 적용하여 사용자 검색 의도에 가장 부합하는 의미해석 결과를 제공하고자 한다. 논문의 온톨로지 지식 베이스 프레임워크 내에서 의미해석 결과를 평가하는데 활용할 수 있는 것은 3가지 정도로 구분할 수 있다. 첫 번째는 검색 대상이면서 의미해석을 제공하는 지식 베이스

이다. 지식 베이스는 도메인 지식을 표현한 저장소이므로 이를 기반으로 의미해석 결과를 평가할 수 있는 모델이 될 수 있다. 두 번째는 검색 시스템에게 있어 가장 중요한 사용자의 검색 이력이다. 더욱 자세하게는 사용자가 입력하는 검색 키워드에 대한 의미해석 결과 중 사용자가 선택한 의미해석 결과이다. 즉, 사용자가 검색문을 입력하고 프레임워크에서 적절한 의미해석 결과를 제공하면, 사용자는 자신의 의도에 가장 부합하는 의미해석 결과를 선택하고 결과를 탐색할 것이다. 이렇게 선택된 질의 의미해석 결과를 저장하여 분석하면 검색에 있어 사용자의 선호도를 획득할 수 있고, 이를 의미해석 평가 모델로 적용할 수 있다. 세 번째는 질의 의미해석 결과의 구조적인 특징을 모델로 적용할 수 있다. 두 개의 검색 키워드로부터 의미해석된 결과는 두 개에서부터 많게는 열 개 이상의 개체로 구성되는 의미해석 결과를 생성할 수 있다. 하지만, 의미해석 결과가 길어질수록 사용자 검색 의도와 맞지 않는 경향이 있다.

5. 사용자 선호 기반 평가 모델

confidence 알고리즘[9]을 활용하여 지식 베이스 기반의 모델과 질의 의미해석 결과 간의 유사도 측정 시에 동일한 패턴의 의미해석 결과는 정확도를 높게 구분할 수 있었다. 하지만, 예를 들어 “영화배우”, “영화감독”의 특성을 동시에 가지고 있는 경우는 confidence 유사도 측정으로는 구분할 수 없다. 뿐만 아니라 지식 베이스 기반의 모델에 존재하지 않는 “영화배우 연출 영화”, “영화배우 각본 영화”의 구분도 불가능하다. 따라서 이러한 단점을 보완하기 위해 사용자의 검색 이력을 추적하고 추적된 검색 이력으로부터 검색 선호도를 추출하고 이를 반영함으로써 좀 더 사용자의 검색의도에 부합하는 의미해석 결과를 제공할 수 있게 된다. 하지만 단순한 검색이력 추적적으로는 특정 인스턴스의 특징을 구분할 수 없게 된다. 예를 들어, “류승완”이라는 사람은 “영화감독”, “영화배우”, “작가”의 특성을 동시에 가지고 있다. 하지만 “류승완”의 경우 “영화감독”으로 연출한 영화가 “영화배우”로 출연한 영화보다 많기 때문에 일반적으로 영화감독으로 더 유명하다. 이러한 정보는 지식 베이스 내에서 “류승완”과 관계된 속성들을 탐색하면 “영화감독”이 “류승완”에게 있어 더 가중치가 높다는 것을 알 수 있다. 또 다른

예로 “유지태”의 경우도 “영화감독”, “영화배우”, “작가”의 특성을 모두 가진다. 하지만 “유지태”의 경우는 “영화배우”로 더 유명하고, 지식 베이스 내의 관계된 속성을 탐색하면 “영화배우”가 “영화감독”보다 가중치가 높다는 것을 알 수 있다. 이러한 예처럼 인스턴스가 다양한 특성을 가지고 있는 경우는 온톨로지 지식 베이스로부터 인스턴스의 특성을 분석함으로써 어떤 특성에 더 높은 가중치를 가지고 있는지 알 수 있다.

사용자 검색 이력은 단순히 사용자가 입력한 검색 키워드를 추적하는 것이 아니다. 사용자가 검색 키워드를 입력하면 시스템에서 검색 키워드의 의미를 해석하여 의미해석 결과를 제공하고, 제공된 의미해석 결과 중 사용자가 검색 의도에 부합한다고 선택한 의미해석 결과를 검색 이력으로 추적한다. 예를 들어, 사용자가 “송강호 영화”라고 입력하면, 시스템은 다수의 의미해석 결과를 제공하고 그 결과 중 “송강호 출연 영화”를 선택하였을 경우 [그림 4]와 같이 의미해석 결과가 히스토리로 추적된다.

logid	timestamp	userid	keywords	interpretation	entities
10003	2012-08-31 09:43:24	home	류승완-연출-영화	[[{"http://www.edn.re.kr/ontologies/semantic_..._75d2033e_6e01_47a5_9563_ea197cd4	
10004	2012-08-31 09:43:28	home	류승완-출연-영화	[[{"http://www.edn.re.kr/ontologies/semantic_..._75d2033e_6e01_47a5_9563_ea197cd4	
10005	2012-08-31 09:43:34	home	류승완-각본-영화	[[{"http://www.edn.re.kr/ontologies/semantic_..._75d2033e_6e01_47a5_9563_ea197cd4	
10013	2012-08-31 09:44:21	home	김현민-연출-영화	[[{"http://www.edn.re.kr/ontologies/semantic_..._75d2033e_6e01_47a5_9563_ea197cd4	
10014	2012-08-31 09:44:30	home	김현민-각본-영화	[[{"http://www.edn.re.kr/ontologies/semantic_..._75d2033e_6e01_47a5_9563_ea197cd4	
10083	2012-08-31 09:51:19	home	송강호-출연-영화	[[{"http://www.edn.re.kr/ontologies/semantic_..._75d2033e_6e01_47a5_9563_ea197cd4	

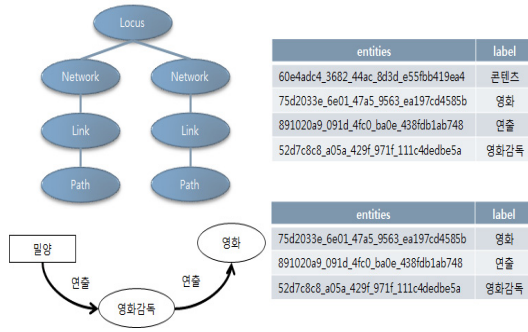
[그림 4] 검색 이력 히스토리

실제 검색 이력 히스토리는 인스턴스의 특성을 추출하여 지식 개체의 집합으로 추상화된 형태로 추적된다.

<표 1> 질의 의미해석 결과의 특성

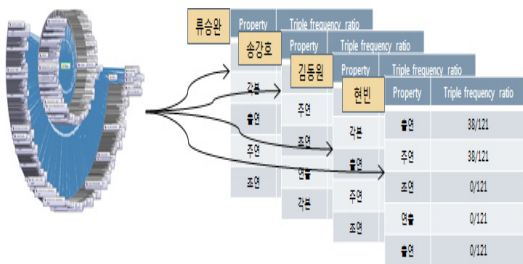
타입	해석 결과	entity	레이블
class	영화	75d2033e_6e01_47a5_9563_ea197cd4585b	영화
Object Property	연출	891020a9_091d_4fc0_ba0e_438fdb1ab748	출연
Instance	류승완	1f587992_20a9_4f95_9ca8_52f0538ebde4	영화배우
		52d7c8c8_a05a_429f_971f_111c4dedbe5a	영화감독

검색 이력의 저장은 질의 의미해석 결과 형태로 저장되며 실제 저장 단위는 질의 의미해석에서 설명한 가장 작은 단위인 패스로 저장한다. 복문의 형태로 구성된 의미해석 결과는 [그림 5]와 같이 패스 단위로 분할하여 검색 이력으로 저장하게 된다.



[그림 5] 패스 단위의 검색 이력 저장

인스턴스의 분석은 [그림 6]과 같이 온톨로지 지식 베이스를 기반으로 이루어진다. 지식 베이스로부터 해당 인스턴스의 모든 트라이플을 탐색하고, 탐색된 트라이플에서 공통 속성을 추출하고 가중치 값을 계산한다. 가중치는 (식 1)와 같이 전체 트라이플 수에 대한 속성의 비율로 정의할 수 있다.



[그림 6] 인스턴스 가중치 분석

$$W_{property} = \left(1 - \frac{\text{count of Triple}_{predicate}(property)}{\text{count of Triple}_{predicate}(properties)}\right) \times \text{normalization Factor}(\min, \max) \quad (\text{식 1})$$

인스턴스 분석 절차를 통해 저장된 검색 이력과 질의 의미해석 결과의 예는 <표 2>와 같다. (식 1)의 기존 confidence 알고리즘을 적용하여 유사도를 측정하면 <표

2>와 같다. 표에서 보는 바와 같이 동일한 유사도가 나타남에 따라 의미해석 결과의 랭킹이 불가능하다. 따라서, 검색 이력과 질의 의미해석 결과 간의 유사도 측정을 위해 (식2)과 같이 modified confidence 알고리즘을 활용한다. modified confidence 알고리즘은 인스턴스 분석에 따라 획득된 가중치를 반영하여 유사도를 측정할 수 있다. <표 3>은 인스턴스의 가중치 분석 값을 적용한 modified confidence 알고리즘의 유사도 측정 결과이다.

$$\text{Modified Confidence}(I, U) = \frac{\sum\{(I \cap U) \cdot L\} / |I \cap U|}{\sum\{(I) \cdot L\} / |I|} \quad (\text{식 2})$$

<표 2> 기존 confidence 유사도 측정 결과

의미해석 결과	검색 이력	유사도	순위
영화배우[1] 연출[1] 영화[1]	영화감독[0.8] 영화배우[0.2] 연출[1] 영화[1]	3/3 = 1.00	1
영화배우[1] 출연[1] 영화[1]	영화감독[0.1] 영화배우[0.9] 출연[1] 영화[1]	3/3 = 1.00	1

<표 3> modified confidence 유사도 측정 결과

의미해석 결과	검색 이력	유사도	순위
영화배우[1] 연출[1] 영화[1]	영화감독[0.8] 영화배우[0.2] 연출[1] 영화[1]	(0.2+1+1)/3 = 0.73	2
영화배우[1] 출연[1] 영화[1]	영화감독[0.1] 영화배우[0.9] 출연[1] 영화[1]	(0.9+1+1)/3 = 0.97	1

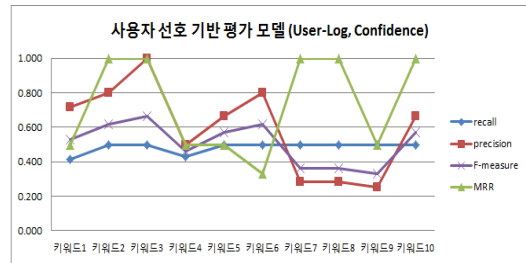
인스턴스의 가중치를 반영하는 modified confidence 알고리즘은 각 요소들을 모두 포함하여야 유사도 측정이 가능하기 때문에 정확률 측면에서 높은 성능을 발휘할 수 있다. 하지만, 요소 중 한 개라도 사용자 검색 이력에 없으면 값을 계산할 수 가 없다. 따라서 검색 이력 상에 없는 형태의 새로운 패턴의 질의 의미해석이 들어오는 경우는 유사도 측정이 불가능하다. 즉, 의미해석 결과 중 동일한 유사도 값을 갖게 되는 후도들이 많아지고, 재현율을 자체가 낮아지게 된다는 단점이 있다.

6. 질의 의미 해석을 위한 사용자 선호 기반 평가 모델의 성능 평가

지식 베이스 기반 평가 모델은 지식 베이스로부터 추출된 패턴으로 나타나는 의미해석 결과에 대해서는 정확률을 높게 평가할 수 있다. 하지만, 재현율이 낮게 나타남에 따라 사용자가 원하는 의미해석 결과가 아닐 경우에 다른 대안을 선택 기회를 제공할 수 없게 된다. 사용자 선호 기반 평가 모델은 지식 베이스를 기반으로 입력된 키워드의 인스턴스를 분석하고, 사용자 로그와 confidence 값을 계산하여 평가함으로써 사용자 개인의 검색 성향에 맞는 의미해석 결과를 제공할 수 있다. <표 4>와 [그림 7]은 사용자 선호 기반 평가 모델의 성능 평가 실험 결과를 정리한 것이다. 실험 결과에서 보는 바와 같이 평균적으로 높은 정확률을 보여주고 있을 뿐만 아니라, 적절한 재현율을 제공하고 있음을 알 수 있다. 특정 검색 패턴에 대해 낮은 정확률을 제공하고 있지만, 사용자 로그가 어떻게 기록되느냐에 따라 편차가 발생할 수 있으며, 평균적으로 높은 정확도를 제공하고 있음을 알 수 있다. 사용자 선호 기반 평가 모델 설계 시에 예측한 대로 지식 베이스 기반 평가 모델에 비해 정확률은 낮지만, 적절한 재현율을 제공함으로써 사용자로부터 피드백을 제공받을 수 있고, 이에 따라 더 낮은 의미해석 결과를 제공할 수 있게 된다.

<표 4> 사용자선호기반평가 모델의 성능 평가 실험 결과

	recall	precision	F-measure	MRR
키워드1	0.417	0.714	0.526	0.500
키워드2	0.500	0.800	0.615	1.000
키워드3	0.500	1.000	0.667	1.000
키워드4	0.429	0.500	0.462	0.500
키워드5	0.500	0.667	0.571	0.500
키워드6	0.500	0.800	0.615	0.333
키워드7	0.500	0.286	0.364	1.000
키워드8	0.500	0.286	0.364	1.000
키워드9	0.500	0.250	0.333	0.500
키워드10	0.500	0.667	0.571	1.000



[그림 7] 사용자 선호 기반 평가 모델의 성능 평가 실험 결과

7. 결론

본 논문에서는 사용자 질의 의미 해석을 위한 선호도 기반 검색 기술을 제안하였다. 제안된 기술에 따라 기존 confidence 유사도 측정 결과와 modified confidence 유사도 측정 결과를 비교하고, 인스턴스의 특성을 추출하여 질의 의미해석 결과의 특성을 분석하여 사용자 선호 기반 평가 모델의 성능 평가 실험 결과를 보면 사용자 선호 기반 평가 모델 설계 시에 예측한 대로 지식 베이스 기반 평가 모델에 비해 정확률은 낮지만, 적절한 재현율을 제공함으로써 사용자로부터 피드백을 제공받을 수 있고, 이에 따라 더 낮은 의미해석 결과를 제공할 수 있게 된다.

향후 연구 과제로 사용자의 선호도에 따른 연관 검색어 추출 기법 및 연관 검색어 유사도를 분석하는 연구를 진행하려 한다.

This research was financially supported by the Ministry of Education, Science Technology (MEST) and National Research Foundation of Korea(NRF) through the Human Resource Training Project for Regional Innovation

참 고 문 헌

- [1] 김정민, 정현숙, “방송 온톨로지 구축 및 매칭 기반의 방송 프로그램 검색”, 한국정보기술학회 제9권 제12호, pp.161-171, 2011년 12월.
- [2] 정휘용, 김경선, 정한민, “시맨틱 검색 기술 동향”, 주간기술동향 통권 1431호, 정보통신산업진흥원, 1432호, pp. 14-27, 2010년 2월.

[3] S. Kang, D. Kim, Y. Lee, S. Hyun, D. Lee, and B. Lee, "A Semantic Service Discovery Network for Large-Scale Ubiquitous Computing Environments," ETRI Journal, vol.29, no.5, 2007.

[4] Andrei Broder, "A taxonomy of web search," ACM SIGIR, volume 36, pp.3-10, 2002.

[5] Rose, D. E., Levinson, D. "In Understanding user goals in Web search," World Wide Web conference (WWW 2004), pp. 13 - 19, 2004.

[6] Kang, I., Kim, G. "In Query type classification for Web document retrieval," ACM SIGIR, pp. 64 - 71, 2003.

[7] S. Park, H. Bae, J. Lee, "End user searching: A Web log analysis of NAVER," Korean Web search engine Library & Information Science Research, pp. 203 - 221, 2005.

[8] Q. Zhou, C. Wang, M. Xiong, H. Wang and Y. Yu, "SPARK: Adapting Keyword Query to Semantic Search," In proceedings of ISWC/ASWC, pp.694-707, 2007.

[9] Burke, R., "Hybrid Recommender Systems: Survey and Experiments," User Modeling and User-Adapted Interaction, Kluwer Academic Publishers, 2002, pp. 331-370.

· 관심분야 : 시맨틱 검색, 온톨로지, 정보검색, 상황인식 컴퓨팅
· E-Mail : leemh@etri.re.kr

도 하 나



· 2013년 2월 : 한남대학교 컴퓨터공학과 컴퓨터공학과(공학사)
· 2013년 3월 ~ 현재: 한남대학교 컴퓨터공학과 석사과정
· 관심분야 : 클라우드 컴퓨팅, 데이터 베이스보안, 시맨틱 검색, 온톨로지
· E-Mail : dohana77@naver.com

최 의 인



· 1982년 : 한남대학교 계산통계학과 (학사)
· 1984년 : 홍익대학교 전자계산학과 (석사)
· 1995년 : 홍익대학교 전자계산학과 (이학박사)
· 1996년 ~ 현재 : 한남대학교 컴퓨터

공학과 교수
· 2003년 : UCLA 방문 교수
· 관심분야 : 시맨틱 웹, 유비쿼터스 컴퓨팅, 모바일, 클라우드 컴퓨팅
· E-Mail : eichoi@hnu.kr

정 훈



· 2012년 2월 : 한남대학교 컴퓨터공학과 컴퓨터공학과(공학사)
· 2012년 3월 ~ 현재: 한남대학교 컴퓨터공학과 석사과정
· 관심분야 : 유비쿼터스, 클라우드 컴퓨팅, 시맨틱 검색, 온톨로지, 정보검색, 상황인식 컴퓨팅

· E-Mail : jeonghoon0322@hanmail.net

이 무 훈



· 2002년 8월 : 한남대학교 컴퓨터공학과 컴퓨터공학과(공학사)
· 2004년 8월 : 한남대학교 컴퓨터공학과(공학석사)
· 2013년 2월 : 한남대학교 컴퓨터공학과(공학박사)
· 2008년 10월 ~ 현재 : 한국전자통신

연구원 차세대스마트TV연구단 선임연구원