

# 사용자 질의 의미 해석을 위한 온톨로지 지식 기반 검색

김난주, 표혜진, 정 훈, 최의인  
한남대학교 컴퓨터공학과

## Ontology Knowledge based Information Retrieval for User Query Interpretation

Nanju Kim, Hyejin Pyo, Hoon Jeong, Euiin Choi

Dept. of Computer Engineering Hannam University Daejeon, Republic of Korea

**요 약** 시맨틱 검색은 논리적으로 표현된 지식 베이스를 사용하여 현재의 키워드 기반 검색보다 더 정확한 결과를 제공할 수 있다. 그러나 일반 사용자는 지식 기반의 복잡하고 정형화된 질의어와 스키마를 잘 알지 못한다. 그래서 검색 시스템은 사용자 키워드의 의미를 해석할 수 있어야 한다. 본 논문에서는 멀티미디어 콘텐츠의 시맨틱 검색을 위한 사용자 질의 의미 해석 시스템을 설명한다. 제안한 시스템은 도메인 온톨로지 기반으로 구축된 지식 베이스의 정형화된 구조에 의미 해석 과정이 통합된 온톨로지 지식 베이스 기반 검색 시스템이다.

**주제어** : 시맨틱 검색, 온톨로지, 지식베이스, 이름 인식, 쿼리 그래프

**Abstract** Semantic search promises to provide more accurate result than present-day keyword matching-based search by using the knowledge base represented logically. But, the ordinary users don't know well the complex formal query language and schema of the knowledge base. So, the system should interpret the meaning of user's keywords. In this paper, we describe a user query interpretation system for the semantic retrieval of multimedia contents. Our system is ontological knowledge base-driven in the sense that the interpretation process is integrated into a unified structure around a knowledge base, which is built on domain ontologies.

**Key Words** : Semantic Search, Ontology, Knowledge Base, Named Entity Recognition, Query Graph Construction

### 1. 서론

시맨틱 검색(semantic search)은 검색 결과의 정확도를 향상시키기 위해 기존의 키워드 기반 정보 검색(information retrieval) 알고리즘 방식을 탈피하여 능동적으로 사용자의 의도를 파악하고, 기존 정보를 가공 분

석하여 정교한 검색 결과를 도출하는 일련의 활동 및 방법론을 통칭한다[1]. 최근에는 시맨틱 검색의 기술 성숙도가 높아짐에 따라 마이크로소프트 Bing, 퀴투라(Quintura), 볼프럼 알파(Wolfram Alpha) 등 국외 검색 서비스에서 시맨틱 검색 기술을 도입하고 있다. 뿐만 아니라, 국내의 네이버, 다음, 네이버와 같은 포털 사이트의

\*본 연구는 미래창조과학부 및 정보통신산업진흥원의 IT/SW창의연구과정의 연구결과로 수행되었음(NIPA-2013-1103)

\*\*본 연구는 교육부와 한국연구재단의 지역혁신인력양성사업으로 수행된 연구결과임

Received 6 March 2014, Revised 5 April 2014

Accepted 20 June 2014

Corresponding Author: Euiin Choi (Hannam University)

Email: eichoi@hnu.kr

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

검색 엔진들도 시맨틱 검색 기술을 도입하고 상용화하기 위한 노력을 기울이고 있다[2].

이처럼, 시맨틱 검색이 정확도 높은 검색을 제공하기 위해서는 첫 번째로 사용자로부터 입력된 부정확한 질의어의 의미를 정확하게 해석하기 위한 방법이 필요하다. 키워드 기반 검색에 익숙한 사용자들은 지식 베이스 기반의 정형화된 질의어와 스키마에 대한 이해가 없기 때문에 기존 키워드 기반의 검색처럼 몇 개의 키워드만으로 검색을 수행하게 된다. 부족하게 입력된 사용자 키워드로부터 사용자 질의의 의미를 해석하는 것은 상당히 어려운 일이다[6, 8]. 두 번째로 잘 구축된 풍부한 지식 베이스가 필요하다. 하지만, 지식 베이스의 구축은 도메인이 한정된다 하더라도 쉽지 않은 일이다. 특히, 실시간 이슈성 키워드와 같이 시간에 따라 변화하는 키워드들을 지식 베이스에 반영하는 것은 더욱 어려운 일이다[3].

따라서 본 논문에서는 정확도 높은 시맨틱 검색을 제공하기 위하여 사용자 검색문의 정확한 의미를 해석하기 위한 온톨로지 기반 사용자 질의 의미 해석 기법을 제안한다. 제안한 질의 의미 해석 기법을 통해 부족한 사용자 검색문으로부터 정확한 검색 의도를 해석할 수 있다.

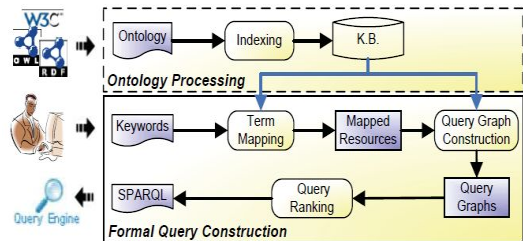
본 논문의 구성은 다음과 같다. 2장에서는 관련연구에 대해 소개하고, 3장에서는 사용자 질의 의미 해석 시스템을 설명하고, 4장에서는 제안 시스템의 검증 및 실험을 설명하고, 마지막으로 5장에서 결론을 맺는다.

## 2. 관련 연구

SPARK 시스템은 사용자 검색 키워드를 지식 베이스 기반으로 의미를 해석하고, 해석된 결과로부터 시맨틱 검색을 위한 정형화된 질의어를 구성하여 검색 결과를 제공하는 시맨틱 검색 시스템이다[4, 5]. [Fig. 1]은 SPARK 프레임워크의 구조를 설명하고 있다. SPARK 시스템은 크게 3가지 단계를 통해 사용자 검색 키워드의 의미를 해석하고, SPARQL(SPARQL Protocol And RDF Query Language)[8]과 같은 정형화된 질의를 생성하여 검색 결과를 제공한다. SPARK 시스템은 키워드 기반 검색에 익숙한 사용자들이 정형화된 질의의 표현 방법을 전혀 모르는 상태에서 일반 키워드 검색과 동일한 검색 키워드를 제공하고 SPARK 시스템이 이를 해석하

여 질의 조합을 구성하는 것이 특징이다.[9]

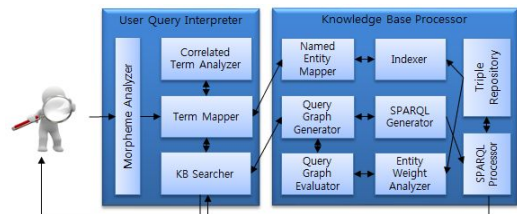
- 개체 식별(term mapping): 사용자 키워드로부터 지식 베이스 내의 지식 개체를 식별
- 질의 그래프 생성(query graph construction): 식별된 지식 개체를 바탕으로 시맨틱 네트워크를 탐색하여 질의 그래프를 구성
- 질의 랭킹(query ranking): 생성된 질의 그래프를 사용자 질의와 지식 베이스를 기반으로 평가



[Fig. 1] SPARK framework

## 3. 사용자 질의 의미 해석 시스템

시맨틱 멀티미디어 콘텐츠 검색을 위한 시스템의 구조는 [Fig. 2]에서 보는 바와 같이 크게 사용자 검색문의 의미 해석기와 지식 베이스 처리기로 구성된다. 지식 베이스 처리기는 사용자 검색 키워드를 지식 베이스 기반으로 개체 식별하고, 식별된 개체로부터 다수의 후보 질의 그래프를 구성한다. 구성된 다수의 질의 그래프를 평가하여 사용자 검색 의도에 가장 부합하는 질의 그래프를 찾고 질의 그래프를 SPARQL로 변환하여 검색 결과를 사용자에게 제공한다. 사용자 검색문의 의미 해석기는 지식 베이스 처리기를 활용하여 사용자 검색문의 의미를 해석한다.



[Fig. 2] The semantic structure of multimedia content retrieval

[Fig. 3]은 사용자 질의 의미 해석 알고리즘을 설명하고 있다. 사용자 질의 해석기는 크게 두 단계를 거치게 된다. 첫 번째 단계는 지식 개체 식별 단계이고, 두 번째 단계는 질의 그래프 구성 단계이다. 개체 식별 단계는 사용자 키워드에 포함된 특정 단어와 지식 베이스 내의 지식 개체와 매핑하는 단계이다. 이 때 지식 개체 식별을 위해 지식 베이스 내의 지식 개체 인덱싱, 사용자 키워드에 대한 형태소 분석, 사용자 키워드 조합 평가 등의 세부 단계들이 필요하다. 그리고 두 번째 단계인 질의 그래프 구성 단계는 식별된 지식 개체로부터 SPARQL을 생성하기 위한 질의 그래프를 구성하는 단계로, 지식 베이스 내의 시맨틱 네트워크 탐색을 통해 해석 가능한 모든 질의 그래프를 구성해야 한다. 해석 가능한 모든 질의 그래프의 구성이 완료되면 지식 베이스를 기반으로 이를 평가하여 사용자 의도에 가장 부합하는 질의 그래프를 선별한다. 이러한 질의 그래프 평가는 지식 베이스 내의 각 개체들 간의 관계와 출현 빈도를 기반으로 가중치 모델을 생성하고 이를 바탕으로 평가하게 된다.

```

1 USER_QUERY_INTERPRETATION(keyword list)
2   user_query_segment_mapping_list =
TERM_MAPPING(keyword list);
3   initialize new empty query_graph;
4   initialize new empty
instance_entity_list;
5   FOR mapping in
user_query_segment_mapping_list
6     add instance_entity_list to mapping;
7     clear instance_entity_list;
8     query_graph_store =
CONSTRUCT_QUERY_GRAPH(mapping);
9     extract shortest query_graph
containing all the
knowledge_entities_of_mapping
from query_graph_store;
10    add shortest_query_graph to
query_graph;
11    IF mapping is not the last one
12      THEN translate
shortest_query_graph into SPARQL query
statement;
13      instance_entity_list =
process SPARQL query with KB;
14    RETURN query_graph;

15 TERM_MAPPING(keyword list)
16 initialize new empty
user_query_segment_mapping_list;
17 FOR keyword in keyword_list
18   initialize new empty mapping_list;
19   find knowledge_entity for keyword;
20   add knowledge_entity to
mapping_list;
21   IF knowledge_entity is class entity
22     THEN add mapping_list to
user_query_segment_mapping_list;
23 RETURN user_query_segment_mapping_list;

24 QUERY_GRAPH_CONSTRUCTION(mapping)
25 initialize new empty query_graph_store;
26 extract class_entity from mapping;
27 FOR knowledge_entity (except
class_entity) in mapping
28   make connections between
knowledge_entity and class_entity;
29   add connections to
query_graph_store;
31 RETURN query_graph_store;
    
```

[Fig. 3] User query semantic analysis algorithms

### 3.1 온톨로지 기반 지식 베이스

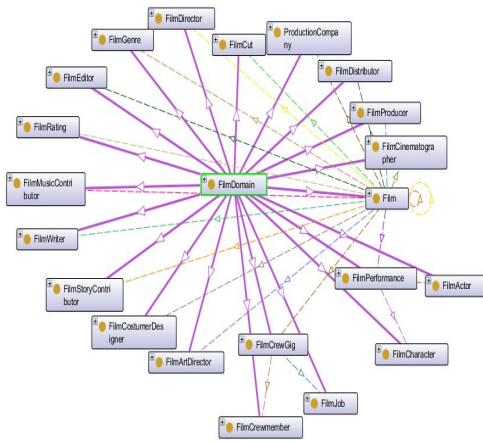
시맨틱 검색에 있어 가장 전제되어야 할 것은 잘 구축된 지식 베이스이다. 제안 시스템에서는 지식 베이스 구축을 위해 프리베이스(freebase)를 참조하여 온톨로지 스키마를 설계하였다. 온톨로지 설계 참조 모델의 선정 기준은 스키마의 지식 표현 범위, 지식 표현 방법, 지식의 풍부성을 고려하였다. 프리베이스의 경우, OWL(Web Ontology Language)[7]을 이용하여 직접 표현 가능하고, 영화, 방송(예능, 드라마), 뮤직비디오 콘텐츠 검색을 위한 프로토타입 시스템의 지식 표현 범위를 충분히 만족하고 있다. 온톨로지 모델 설계를 위해 BBC 프로그램, IMDb(Internet Movie Database) 등의 온톨로지를 분석하였으나, 특정 도메인의 지식만을 표현하기 위해 설계되어 있으며, 상이한 표현 방법을 사용하고 있기 때문에 이들의 통합이 어렵다. 프리베이스에서는 타입(type)을 특정 토픽에 대한 is-a 관계로 나타내며, 그 종류는 열거형 타입(enumerated type), 상속형 타입(included type), 다중 필드 표현을 지원하는 CVT(Compound Value Type)으로 구성된다. 프리베이스에서 속성은 쌍으로 연결된 두 개의 타입의 관계를 나타내며, 각 속성의 아크는 항상 단방향성 아크를 의미한다. 아크의 네이밍 규칙은 아크가 속한 영역의 ID, 아크가 시작되는 쪽의 클래스 아이디어에 아크(속성)의 이름을 붙임으로써 아크 ID의 유일성을 유지한다.

참조 모델을 바탕으로 구축된 영화, 방송, 뮤직 비디오를 도메인으로 하는 지식 베이스의 구성은 <Table 1>과 같다.

<Table 1> The configuration of the knowledge base

	Class	Object Property	Datatype Property	Instance
Entity	84	83	51	2,123

[Fig. 4]는 영화 도메인을 구성하는 각 클래스를 보여주고 있으며, <Table 2>는 영화 콘텐츠를 표현하는 클래스와 레이블의 예를 보여주고 있다.



[Fig. 4] A subclass of the film domain

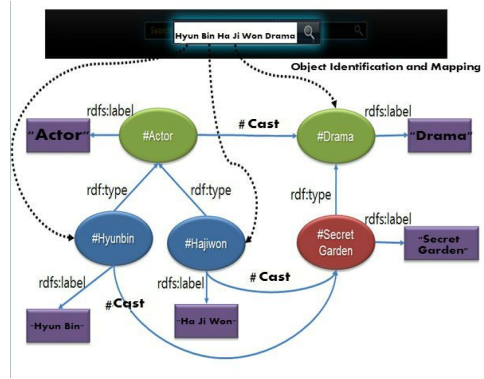
<Table 2> Movie content classes and tables

Class	Label	Class	Label
Film	Movie	FilmDistributor	Film distributors
FilmActor	Film actor	FilmEditor	Film editor
FilmFestival	Film	FilmGenre	Movie Genres
FilmArtDirector	Art Director	FilmMusicContributor	Film music producer
FilmCharacter	Movie Cast	FilmPerformance	Movies starring Events
FilmCostumeDesigner	Film Costume Designer	FilmProducer	Moviemaker
FilmCrewGig	Film crew event	FilmRating	Movie Ratings
FilmCrewmember	Film crew	FilmWriter	Screenplay
FilmCut	Film Editing Edition	FilmStoryContributor	Film author
FilmDirector	Film Director	ProductionCompany	Film production

### 3.2 개체 식별

개체 식별 과정은 사용자로부터 입력받은 검색 키워드를 분석하여 지식 베이스 내에서 존재하는 지식 개체를 찾는 과정이다. [Fig. 5]는 개체 식별 및 매핑 과정을 설명하고 있다. 사용자가 “현빈 하지원 드라마”라는 검색 키워드를 입력하게 되면, 시스템은 형태소 분석을 통해 입력된 키워드를 분리하고 “현빈”, “하지원”, “드라마”의 각 키워드에 해당하는 지식 개체를 지식 베이스로부터

식별한다. 개체 식별 결과로는 “현빈”의 경우 배우 “#Hyunbin”이라는 인스턴스로 식별되고, “하지원”은 배우 “#Hajiwon”이라는 인스턴스로 식별된다. 그리고 “드라마”는 클래스 “#Drama”로 식별된다.



[Fig. 5] Individual identification

이러한 개체 식별을 위해 시스템은 지식 베이스 내의 각 지식 개체에 대한 레이블을 중심으로 인덱싱 과정을 수행한다. 레이블 및 지식 개체 인덱싱을 위해 시스템은 루씬(Lucene)을 활용하였다. 또한 사용자가 입력한 검색 키워드로부터 의미 있는 형태소를 추출하기 위해 KAIST에서 개발한 한나눔 형태소 분석기를 활용하였다. 형태소 분석기를 통해 조사와 어미를 제외한 모든 품사를 활용한다.

<Table 3>에서와 같이 사용자가 입력한 “현빈 하지원 드라마”를 형태소 분석하여 명사를 추출하고 추출된 키워드의 수를 N개로 하였을 때 2(N-1)개로 구성되는 키워드 조합으로부터 TF/IDF(Term Frequency/ Inverse Document Frequency)와 사용자 키워드 순서 관계로부터 최적 조합을 계산한다. 가장 높은 점수의 최적 조합으로부터 개체 식별 결과를 생성한다.

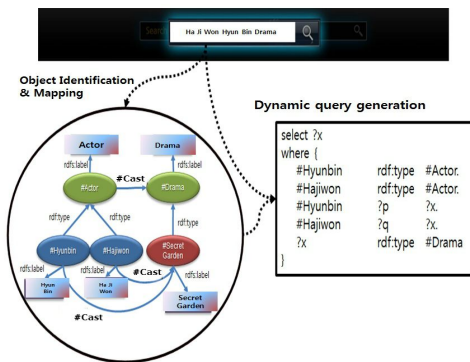
<Table 3> Optimal combination of user search keywords

```
Original Query String: Hyun Bin Ha Ji Won Drama
Result of Sentence "Hyun Bin Ha Ji Won Drama" Analyze by HanNanum:
Hyun Bin-ncn, Ha Ji Won-ncn, Drama-ncn,
Partition #0 of Query: {{Hyun Bin, Ha Ji Won, Drama}}
Partition #0 of Query: {{Hyun Bin}, {Ha Ji Won, Drama}}
Partition #0 of Query: {{Hyun Bin, Ha Ji Won}, {Drama}}
Partition #0 of Query: {{Hyun Bin}, {Ha Ji Won}, {Drama}}
Best Partition: {{Hyun Bin}, {Ha Ji Won}, {Drama}},0.9882412354151408
```

### 3.3 질의 그래프 생성 및 평가

개체 식별이 완료되면, 지식 베이스로부터 식별된 개체를 모두 포함하는 부분 그래프를 찾는 질의 그래프 생성 과정을 수행한다. 부분 그래프를 찾는 과정에서 식별된 지식 개체를 포함하는 부분 그래프는 다양하게 존재 가능하다. 이러한 부분 그래프는 사용자가 입력한 검색 키워드로부터 해석 가능한 다양한 후보 해석 대안이 될 수 있으며, 이러한 후보 해석 대안을 지식 베이스 기반으로 평가하여 가장 적합한 해석 대안을 결정하고, 결정된 해석 대안으로부터 SPARQL 형태의 질의 문장을 생성한다.

[Fig. 6]에서 보는 바와 같이 “현빈 하지원 드라마”에 대한 지식 개체 식별이 완료되면, #Hyunbin, #Hajiwon으로부터 #Actor 클래스를 찾고 #Actor 클래스로부터 #Drama로 연결되는 부분 그래프를 탐색한다. 부분 그래프를 탐색 시에는 MST(Minimum Spanning Tree)와 A\* 알고리즘을 사용한다[7, 9]. 이 때 동일한 토폴로지(topology) 내에서 #Actor로부터 #Drama로 연결되는 그래프 경로는 중간에 다양한 ObjectProperty로 연결될 수 있다.



[Fig. 6] Query semantic analysis procedures

[Fig. 6]에서는 #출연하다 라는 ObjectProperty로만 연결이 되어 있지만, #연출하다, #제작하다 등의 다양한 ObjectProperty가 존재할 수 있으며 이들이 후보 해석 대안이 된다. 이러한 후보 해석 대안에 대한 랭킹을 위한 논문에서는 지식 개체 가중치 행렬(entity weight matrix)를 구축하고 이를 기반으로 후보 해석 대안을 평가한다. 가중치 행렬의 구축은 초기 온톨로지 적재 시 온톨로지 상의 모든 개체(class, instance, property)에 대한

출현 빈도를 통해 가중치 모델을 생성한다. 각 개체의 가중치 계산식은 아래와 같다. 가중치의 최댓값(0), 최솟값(1)을 이용하여 각 가중치를 정규화하고 이를 가중치 모델로 생성한다. 따라서 현빈과 하지원의 경우, 지식 베이스 상에 배우로써 드라마에 출연하는 빈도가 높기 때문에 “현빈 하지원이 출연한 드라마”를 최적 해석 결과로 사용자에게 검색 결과와 함께 제공하게 된다.

식 (1) 클래스에 대한 가중치를 구하는 식이며, 획득된 전체 인스턴스의 개수에 대한 각 클래스의 인스턴스 개수의 비율로 표현된다. 식 (2)는 프로퍼티에 대한 가중치를 구하는 식이며, 전체 프로퍼티를 predicate으로 갖는 개수에 대한 각 프로퍼티를 predicate으로 갖는 개수의 비율로 표현된다. 식 (3)은 인스턴스에 대한 가중치를 구하는 식이며, 전체 트리플 개수에 대한 각 인스턴스를 domain이나 range로 갖는 트리플의 개수의 비율로 표현된다.

$$W_{class} = \left(1 - \frac{\text{count of Instance(class)}}{\text{count of Instance(classes)}}\right) \times \text{normalizationFactor}(\min, \max) \quad (\text{eq. 1})$$

$$W_{property} = \left(1 - \frac{\text{count of Triple}_{predicate}(property)}{\text{count of Triple}_{predicate}(properties)}\right) \times \text{normalizationFactor}(\min, \max) \quad (\text{eq. 2})$$

$$W_{instance} = \left(1 - \frac{\text{count of Triple(instance)}}{\text{count of Triple(instances)}}\right) \times \text{normalizationFactor}(\min, \max) \quad (\text{eq. 3})$$

## 4. 시스템 검증 및 실험

### 4.1 실험환경 구성

제한된 검색 시스템의 검증을 위해 10개의 검색 예제를 정하였다. 그리고 재현율(recall rate)과 정확률(precision rate) 평가를 위해 영화, 드라마, 예능 콘텐츠 기준으로 각 검색 예제 별로 10개의 적합 콘텐츠 집합을 선별하고, 10개의 부적합 콘텐츠 집합을 선별하였다. 또한 각 예제의 정답 집합으로부터 MRR(Mean Reciprocal Rank) 평가를 실험하였다. 예제별로 TF/IDF 기반의 키워드 검색 방법과

제안 시스템에서 검색을 수행하여 검색 성능을 평가하였다. 각 평가 방법의 정의는 아래와 같다.

$$recall = \frac{|{\{relevant\ documents\}} \cap {\{retrieved\ documents\}}|}{|{\{retrieved\ documents\}}|}$$

(eq. 4)

$$precision = \frac{|{\{relevant\ documents\}} \cap {\{retrieved\ documents\}}|}{|{\{relevant\ documents\}}|}$$

(eq. 5)

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank}$$

(eq. 6)

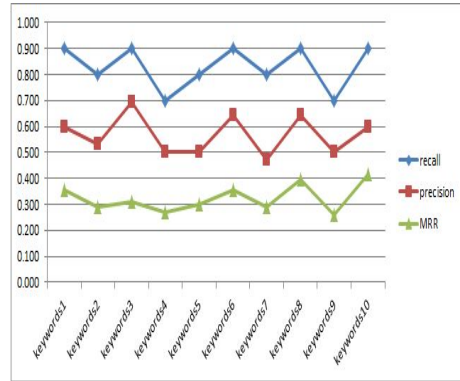
### 4.2 실험 결과

<Table 4, 5>와 [Fig. 7, 8]은 각 시스템에 대하여 검색 예제별로 평가한 재현율, 정확률, MRR을 설명하고 있다. 그리고 [Fig. 7]은 각 시스템에 대하여 검색 예제로 평가한 재현율, 정확률, MRR의 평균값을 설명하고 있다. 실험 결과에서 보는 바와 같이 TF/IDF 기반 검색은 다른 시스템에 비하여 재현율은 높으나 정확률이 급격히 감소하고 있음을 알 수 있다. 그리고 기법의 특성상 타 시스템에 비하여 재현율이 높다. 더욱이 TF/IDF 기반 시스템의 경우 MRR이 재현율에 비해 현격히 낮은 성능을 보여주는 것을 알 수 있다. 검색 결과에는 포함되어 있지만, 검색 결과가 나타나는 순위가 낮음을 보여주는 것이다.

논문에서 제안하는 사용자 질의 확장 검색의 경우, 기존 시스템과의 객관적인 성능 평가 및 비교는 불가능하다. 하지만, 제안 시스템에서는 질의 확장기를 통하여 정확한 검색 결과를 제공한다.

<Table 4> TF / IDF Results

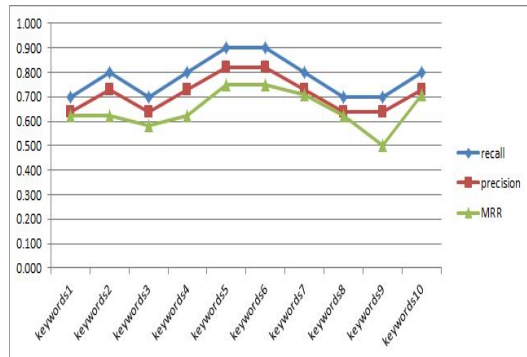
	key word s1	key word s2	key word s3	key word s4	key word s5	key word s6	key word s7	key word s8	key word s9	key word s10
recall	0.90	0.80	0.90	0.70	0.80	0.90	0.80	0.90	0.70	0.90
precision	0.60	0.53	0.69	0.50	0.50	0.64	0.47	0.64	0.50	0.60
MRR	0.35	0.29	0.31	0.26	0.30	0.35	0.29	0.39	0.25	0.41
	4	2	3	9	0	4	2	6	8	7



[Fig. 7] Experimental results for a keyword search (TF / IDF)

<Table 5> Experimental results of the proposed system

	key word s1	key word s2	key word s3	key word s4	key word s5	key word s6	key word s7	key word s8	key word s9	key word s10
recall	0.70	0.80	0.70	0.80	0.90	0.90	0.80	0.70	0.70	0.80
precision	0.63	0.72	0.63	0.72	0.81	0.81	0.72	0.63	0.63	0.72
MRR	0.62	0.62	0.58	0.62	0.75	0.75	0.70	0.62	0.50	0.70
	5	5	3	5	0	0	8	5	0	8



[Fig. 8] Experimental results for a keyword search (proposed system)

## 5. 결론

본 논문에서는 온톨로지 지식 베이스를 기반으로 사용자 질의 의미 해석과 확장 기법을 제안하였다. 사용자

질의 확장 검색의 경우, 기존 시스템과의 객관적인 성능 평가, 비교 분석은 불가능하다. 하지만, 제안된 기법에 따라 프로토타입을 구현하고 실험을 통하여 TF/IDF 기반의 키워드 기반 검색과 타 시맨틱 검색 시스템보다 정확한 검색 결과를 제공하는 것을 확인 할 수 있었다. 제안된 시스템은 한정된 도메인에 대한 지식 베이스를 구축하고 프로토타입 형태로 검색 결과 실험을 수행하였으나, 검색 성능에 있어 타 시스템과 비교하여 우수한 성능 즉, 정확한 검색 결과를 제공하고 있다.

향후 연구 과제로 사용자 검색 패턴 분석을 통해 사용자 선호 기반 질의 그래프 평가 기법에 대한 연구를 진행할 것이다.

### ACKNOWLEDGMENTS

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.

"This research was supported by the MSIP(Ministry of Science, ICT & Future Planning), Korea, under IT/SW Creative research program supervised by the NIPA(National IT Industry Promotion Agency)" (NIPA-2013-1103)

### REFERENCES

[1] JM, Kim, HS, Jeong, "Ontology-based TV program search and matching", Korea Institute of Information Technology, Vol. 9, No. 12, pp. 161-171, 2011.

[2] HW, Jeong, KS, Kim, HM, Jeong, "Semantic Search Technology Trends", Weekly Technology Trends 1431, Industry Promotion Agency, Vol. 1432, pp. 14-27, 2010.

[3] DK, Lee, JH, Kwon, "Considering your interest in recent social search algorithm", Korea Institute of

Information Technology, Vol. 9, No. 4, pp. 187-194, 2011.

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

[5] T. Tran, P. Cimiano, S. Rudolph and R. Studer, "Ontology-Based Interpretation of Keywords for Semantic Search", LNCS Vol. 4825, 2007.

[6] E. Mäkelä, "Survey of Semantic Search Research", in Proceedings of the Seminar on Knowledge Management on the Semantic Web, 2005.

[7] OWL: Web Ontology Language Reference (<http://www.w3c.org/TR/owl-ref/>), W3C, 2004.

[8] Y. Lei, V. Uren and E. Motta, "SemSearch: A Search Engine for the Semantic Web", LNCS(LNAI), Vol. 4248, 2006.

[9] M. Lee, M. Kim, J. Yeom, K. Lee, Y.suh, H. kim and J. cho, "Ontological Knowledge Base-Driven Framework for Semantic Multimedia Contents Retrieval", Proc. 14<sup>th</sup> Int. Conf. Advanced Communication Technology, 2012.

#### 김 난 주(Kim, Nanju)



- 2014년 2월 : 한남대학교 컴퓨터공학과(공학사)
- 2014년 3월 ~ 현재 : 한남대학교 컴퓨터공학과(석사과정 재학중)
- 관심분야 : 시맨틱, 클라우드 컴퓨팅, 빅데이터
- E-Mail : 91knj@naver.com

#### 표 혜 진(Pyo, Hyejin)



- 2014년 2월 : 한남대학교 컴퓨터공학과(공학사)
- 2014년 3월 ~ 현재 : 한남대학교 컴퓨터공학과(석사과정 재학중)
- 관심분야 : 시맨틱, 클라우드 컴퓨팅, 빅데이터
- E-Mail : ninanaena@gmail.com

정 훈(Jeong, Hoon)



- 2012년 2월 : 한남대학교 컴퓨터공학과(공학사)
- 2014년 2월 : 한남대학교 컴퓨터공학과(공학석사)
- 2014년 3월 ~ 현재 : 한남대학교 컴퓨터공학과(박사과정 재학중)
- 관심분야 : 클라우드 컴퓨팅, 상황 인식 컴퓨팅, 빅데이터

· E-Mail : jeonghoon0322@hanmail.net

최 의 인(Choi, Euiin)



- 1982년 2월 : 한남대학교 계산통계학과(학사)
- 1984년 8월 : 홍익대학교 전자계산학과(석사)
- 1995년 8월 : 홍익대학교 전자계산학과(이학박사)
- 1996년 3월 ~ 현재 : 한남대학교 컴퓨터공학과 교수

· 2003년 3월 : UCLA 방문 교수

· 관심분야 : 시맨틱 웹, 유비쿼터스 컴퓨팅, 모바일, 클라우드 컴퓨팅, 빅데이터

· E-Mail : eichoi@hnu.kr