

효율적 자원 탐색을 위한 소셜 웹 태그들을 이용한 동적 가상 온톨로지 생성 연구*

이현정

연세대학교 공과대학 글로벌융합공학부
(hjlee5249@gmail.com)

손미애

성균관대학교 공과대학 시스템경영공학과
(myesohn@skku.edu)

본 논문에서는 네트워크 기반 대용량의 자원들을 효율적으로 검색하기 위해 사용자의 요구사항에 기반해 검색에 요구되는 태그들 간의 의미론에 기반한 동적 가상 온톨로지 (Dynamic Virtual Ontology using Tags: DyVOT)를 추출하고 이를 이용한 동적 검색 방법론을 제안한다. 태그는 소셜네트워크 서비스를 지원하거나 이로부터 생성되는 정형 및 비정형의 다양한 자원들에 대한 자원을 대표하는 특성을 포함하는 메타적 정보들로 구성된다. 따라서 본 연구에서는 이러한 태그들을 이용해 자원의 관계를 정의하고 이를 검색 등에 활용하고자 한다. 관계 등의 정의를 위해 태그들의 속성을 정의하는 것이 요구되며, 이를 위해 태그에 연결된 자원들을 이용하였다. 즉, 태그가 어떠한 자원들을 대표하고 있는지를 추출하여 태그의 성격을 정의하고자 하였고, 태그를 포함하는 자원들이 무엇인지에 의해 태그간의 의미론적인 관계의 설정도 가능하다고 보았다. 즉, 본 연구에서 제안하는 검색 등의 활용을 목적으로 하는 DyVOT는 태그에 연결된 자원에 근거해 태그들 간의 의미론적 관계를 추출하고 이에 기반하여 가상 동적 온톨로지를 추출한다. 생성된 DyVOT는 대용량의 데이터 처리를 위해 대표적인 예로 검색에 활용될 수 있으며, 태그들 간의 의미적 관계에 기반해 검색 자원의 뷰를 효과적으로 좁혀나가 효율적으로 자원을 탐색하는 것을 가능하도록 한다. 이를 위해 태그들 간의 상하 계층관계가 이미 정의된 시맨틱 태그 클라우드인 정적 온톨로지를 이용한다. 이에 더해, 태그들 간의 연관관계를 정의하고 이에 동적으로 온톨로지를 정의하여 자원 검색을 위한 동적 가상 온톨로지 DyVOT를 생성한다. DyVOT 생성은 먼저 정적온톨로지로부터 사용자 요구사항을 포함하는 태그를 포함한 부분-온톨로지들을 추출하고, 이들이 공유하는 자원의 정도에 따라 부분-온톨로지들 간의 새로운 연관관계 여부를 결정하여 검색에 요구되는 최소한의 동적 가상 온톨로지를 구축한다. 즉, 태그들이 공유하는 자원이 무엇인가에 의해 연관관계가 높은 태그들 간에는 이들의 관계를 설명하는 새로운 클래스를 가진 생성된 동적 가상 온톨로지를 이용하여 검색에 활용한다. 온톨로지의 인스턴스는 자원으로 정의되고, 즉 이는 사용자가 검색하고자 하는 해로서 정의된다. 태그들 간의 관계에 의해 생성된 DyVOT를 이용해 기존 정적 온톨로지나 키워드 기반 탐색에 비해 검색해야 할 자원의 양을 줄여 검색의 정확성과 신속성을 향상 시킨다.

논문접수일 : 2012년 08월 23일 논문수정일 : 2013년 01월 25일 게재확정일 : 2013년 02월 12일

투고유형 : 국문일반 교신저자 : 손미애

1. 서론

웹 2.0 및 3.0에 기반한 정보시스템의 대중적 보

급은 대용량의 정보생성을 야기했다. 이는 정보의
량을 풍부하게 하는데 큰 기여를 하였으나, 사용
자의 사용 목적에 맞는 양질의 가치 있는 정보의

* 본 연구는 지식경제부 및 정보통신산업진흥원의 “IT명품인재양성사업”의 연구결과로 수행되었음(NIPA-2012-H0201-12-1001), 그리고 방위사업청과 국방과학연구소의 지원으로 수행되었음(UD110006MD).

신속한 추출을 위해서는 보다 많은 노력을 필요로 하게 하였다. 즉, 필요에 맞는 정보 자원 획득을 위한 양질의 정보 분석 및 정제에 대한 요구를 증대시키고 있다. 즉, 온라인상에서 실시간으로 생성되는 무한대의 정보들로부터 사용자가 필요로 하는 정보를 신속 정확히 추출하는 것은 매우 중요한 일이다.

이와 같이 대용량의 데이터 처리를 위해 구글의 맵리듀스 등은 정보의 분류를 위한 정보 지도를 생성하고 이를 활용하고 있다(Yoo, 2012). 그러나, 이의 한계점은 정보 분류에 있어 의미적 관계의 정의가 쉽지 않다는 점이다(Park et al., 2011). 또한 기존의 대부분의 검색이 키워드에 기반하여 이루어지고 있으나, 이들 간의 의미적 관계를 고려한 검색에 대한 논의는 부족하다. 따라서, 본 논문에서는 대용량의 데이터 처리의 효율성을 위해 기존 키워드방식과 달리, 키워드를 태그에 대응하여 태그들간의 의미적 관계를 활용하여 검색의 부를 좁혀나가 효율적 탐색이 가능한 동적 가상 온톨로지 DyVOT을 제안한다.

본 논문에서 제안하는 의미론적 관계에 기반한 동적 카탈로그는 사용자 A와 사용자 B가 입력한 특정 키워드가 같더라도 동반 입력된 키워드들이 무엇인지에 따라, 즉 사용자의 검색 목적에 따라 온톨로지 상의 다른 계통 경로를 따라 해로서의 자원의 추출이 가능하도록 지원한다. 또한 동반 키워드 간의 관계가 온톨로지 상의 같은 계통인 경우 검색할 키워드의 수를 줄여 검색의 효율을 높일 수 있다. 즉, 사용자의 요구사항에 따라 동적으로 데이터 카탈로그의 가상적 변형이 가능하도록 하여 정보 탐색의 비용을 줄이고 사용자의 목적에 부합한 정보를 추출하며, 검색의 효율성도 높이도록 한다. 따라서, 본 연구는 의미론적 관계에 기반한 정보 검색에 따라 해 검색의 신속성,

정확성 및 효율성의 향상을 목적으로 한다.

본 연구에서 제안된 동적 가상 온톨로지(Dynamic Virtual Ontology using Tags : DyVOT)는 사용자가 원하는 웹 자원의 효율적 탐색을 위해 자원의 메타 데이터인 태그를 이용한다. 태그란 소셜 태깅 시스템들의 사용자들이 자원(웹페이지, 이미지, 동영상, 단문 등)에 일련의 단어들로 주석을 붙인 것을 말한다(Specia and Motts, 2007). 따라서, 태그는 자원의 특성과 정보 등을 반영한다고 볼 수 있다. 사용자의 사용 목적에 맞는 해를 찾기 위해 시시각각으로 생성되는 대량의 정보들을 일일이 분석하기란 쉽지 않은 상황에서, 다행히 네트워크 기반 웹(Web), 클라우드 컴퓨팅이나 소셜네트워크 서비스(SNS) 등에 기반해 작성된 정보 또는 문서 등은 대부분 태그(Tag)를 포함하고 있다. 따라서 본 연구에서 제안하는 DyVOT는 태그 클라우드에 태그들 간의 의미적 계층 관계(Semantic relationships)를 정의한 온톨로지 기반 시맨틱 클라우드(Semantic Cloud)를 이용하여 사용자의 요구사항에 따라 검색의 효율성을 향상시킬 수 있는 동적 가상 온톨로지로 작성된다. 온톨로지 기반 시맨틱 클라우드의 태그분류는 의미상의 계층관계를 정의한 트리구조로 구성되어 있고(<http://semanticcloud.sandra-siegel.de/>), 이를 본 연구에서는 정적온톨로지라 한다. 이로부터 사용자 요구사항에 대응되는 태그 조합에 따라 생성되는 개념적 연관관계를 동적으로 정의한 가상의 온톨로지 DyVOT는 정적온톨로지로부터 추출된 부분-온톨로지들의 통합으로 생성된다. 태그 조합 중 정적온톨로지에서도 같은 경로를 따라 존재하는 태그들을 계통-태그라 하고, 이들 태그 중 가장 상위에 위치한 태그를 탐노드로 하고, 이의 하위 클래스들과 자원으로서의 인스턴스를 모두 포함한 부분-온톨로지를 추출한다. 태그들의 계통에

따라 복수개의 부분-온톨로지들을 추출할 수 있다. 이는 태그를 계통에 따라 그룹을 구성함으로써 실제 검색에 사용되는 태그들은 부분-온톨로지의 최하단에 위치한 태그들만으로 구성됨으로 실제 검색에 사용될 태그의 수를 온톨로지를 사용하여 줄여 검색의 부를 줄이는 효과가 있다. 즉, 검색의 대상이 되는 자원의 수를 줄여 검색의 효율성을 높이는 효과가 있다.

DyVOT 생성을 위해 태그들 간의 계층관계 이외의 연관 관계 정의가 요구된다. 이를 위해 태그들 간의 공유하는 자원의 정도를 이용한다. 즉, 비교되는 태그들 간의 공유자원이 많을수록 그들 간의 개념적 연관관계가 높다 할 수 있다. 따라서 이들을 이용해 새로운 클래스를 정의하여 부분-온톨로지간의 새로운 관계를 형성하여 동적 가상 온톨로지 DyVOT를 구축한다.

태그 기반 데이터 카탈로그 작성 연구에 따른 기대 효과는 다음과 같다. 태그 클라우드에 기반한 의미론적 데이터 카탈로그는 데이터 분류체계로서 데이터의 효율적 탐색 및 저장을 지원하는 것이 가능하며 동적 데이터 카탈로그를 이용해 사용 목적에 부합한 동적 데이터의 탐색을 효율적으로 지원한다.

본 논문은 제 2장에서 관련문헌연구를 살펴보고, 제 3장에서는 DyVOT 생성 프로세스에 대해 살펴본다. 제 4장에서는 부분-온톨로지 생성을 논의하고, 제 4장에서는 태그들 간의 계층 및 연관관계를 정의하고, 제 5장에서 생성된 부분-온톨로지들 간의 관계 생성을 논의한다. 제 6장에서는 부분-온톨로지를 이용해 DyVOT의 생성을 설명한다. 제 7장에서는 활용예제를 제시하고, DyVOT의 우수성을 논의한다. 마지막으로 결론 및 향후 연구 방향에 대해 논의하기로 한다.

2. 관련 문헌연구

대용량의 데이터 처리의 가능성은 컴퓨터 프로세싱 성능의 향상과 밀접한 관련이 있다(McKinsey Global Institute, 2011). 그러나 단순히 컴퓨터 성능의 향상이 대용량의 데이터 처리의 정확성과 신속성을 보장하지는 않는다(Carter, 2011). 따라서 보다 신속하고 정확한 데이터 프로세싱을 위한 다양한 연구가 이루어져 왔다(Zikopoulos, 2012). 콘텐츠에 기반한 탐색이 일반적으로 가능하겠으나(Lv et al., 2006; Arias and Sánchez, 2003; Ganguly et al., 2006; Yang, 2007) 실제로 웹상의 정보자원은 콘텐츠를 분석하여 필요로 하는 정보자원을 추출하기에는 너무 많으며 또한 다양한 형태의 자원들이 산재해 있다(Dean and Ghemawat, 2004; Dean and Ghemawat, 2008; Helvey, 2007; Madden, 2002). 따라서 정보자원의 효율적 추출을 위한 정보 요약과 정보 추출 방법론에 대한 연구가 요구되고 있다(Yoo, 2012). 이를 위해 태그와 태그 클라우드 및 태그를 기반으로 한 자원 클러스터링에 관한 많은 연구가 다음과 같이 진행되었다.

2.1 태그 클라우드

태그 클라우드는 메타 데이터에서 얻어진 태그들을 분석하여 중요도 또는 인기도 등을 고려하여 시각적으로 늘어놓으며, 태그의 색상, 크기 및 굵기 등으로 중요도를 달리 표현한다. 그러나 이들 태그들 간의 의미적 관계에 대한 정의는 표현되어 있지 않다. 따라서, 태그들간의 내재된 연관관계를 추론하여 태그 클라우드를 생성한 시맨틱 태그 클라우드(Semantic tag-cloud)에 관한 연구들이 등장하게 되었다(<http://semanticcloud.sandra-siegel.de/>). 이들 시맨틱 태그 클라우드는 트리구조 상에서의

계층적 관계 등을 표현하고 있다. 따라서 기존의 태그 클라우드에서 태그들 간의 시맨틱 기반 계층 관계를 고려한 형태의 클라우드라고 볼 수 있다. 즉, 현재 작성된 태그 클라우드는 의미론적 관계의 고려 없이 구성되었거나 의미론적 관계가 고려되었더라도(Angeletou and Sabou, 2007; Hotho, 2006a; Zhang et al., 2006; Schrammel et al., 2009; Jung, 2011; Jung, 2012) 태그 클라우드의 태그들 간의 의미적 계층 관계만 고려한 상황이다.

2.2 태그 클러스터링

대용량의 데이터로부터 사용자가 원하는 정보를 검색하기란 사막에서 바늘을 찾는 것만큼이나 어려운 일이 되었다. 비록 키워드에 기반한 검색이 범용으로 이루어지고 있으나, 단순 키워드에 기반한 검색으로 대용량의 정보자원으로부터 신속 정확하게 사용자가 원하는 정보를 탐색하기란 여전히 용의하지 않다. 따라서 사용자의 요구사항에 맞는 정보 해를 효과적으로 찾기 위해(Park et al., 2011) 데이터 클러스터링과 관련된 연구들이 이루어져 왔다(Hassan-Montero and Herrero-Solana, 2007; Jaschke 2006 Knautz et al., 2010; Salonen, 2007; Yoo et al., 2011). 그러나, 네트워크상의 정보 자원의 콘텐츠를 모두 분석하여 클러스터링을 한다는 것은 하루에도 몇 백만 건의 문서 또는 정보 들이 창출되는 상황에서는 거의 불가능하다. 따라서 메타데이터인 태그를 주로 이용하는 연구들이 이루어져 왔다(Zikopoulos, 2012 Yoo, 2011). 그러나 현재 태그를 이용한 연구들은 태그 기반의 정보 분류 방법론에만 초점을 맞추어 이루어져 왔다. 또한 태그는 단어로 표현되므로 이음동의어와 동음이의어 등이 존재 할 수 있으며, 이에 의미에 기반한 이음동의어 등의 단어분류(Specia and Motts, 2007,

Jung, 2011; Jung, 2012) 및 단어들 간의 의미기반 상위/하위 등의 계층 관계에 기반한 정보 분류체계 등의 연구 들이(Angeletou and Sabou, 2007; Hotho, 2006a; Zhang et al., 2006; Schrammel et al., 2009) 진행되어 왔다. 그러나 이들 연구들의 문제점은 태그들 간의 의미적 개념 분류에만 초점을 두었을 뿐 개념적 연관관계를 추출하고 이에 기반해 검색에 활용하는 데에는 한계가 있다(Knautz et al., 2010). 또한 태그들 간의 의미적 관계에 기반하여 사용자가 검색하고자 하는 키워드 간의 클러스터링이나, 클러스터에 기반하여 검색의 뷰를 줄이고자 하는 노력은 아직 미흡한 상황이다.

따라서 본 연구에서는 제안하는 DyVOT은 시맨틱 태그 클라우드를 태그 기반 정적 온톨로지로 하여, 이에 정의된 계층관계 이외에 태그들 간의 연관관계도 고려한 형태의 동적 가상 온톨로지를 제안한다. DyVOT은 키워드간의 연관관계를 고려하여 검색대상의 뷰를 줄여 효과적 검색이 가능하도록 한다. 제안하는 DyVOT에 대해서 다음 장에서 살펴 보도록 하자.

3. DyVOT 생성 프로세스

정적 데이터 카탈로그(Dean and Ghemawat, 2004; 2008)가 사용자의 요구사항에 따라 정보 자원을 찾기 위해 항상 정적인 탐색을 허용하는 것과 달리 DyVOT은 사용자 요구사항에 매칭되는 계통-태그들에 따라 검색된 복수개의 부분-온톨로지들을 이용해 병렬 검색 처리를 통해 공유된 자원을 동적으로 검색하여 검색의 신속성과 정확성을 높인다.

부분-온톨로지 추출은 사용자의 요구사항에 기반한다. 요구사항은 태그에 대응이 되고, 태그기반 정적온톨로지로부터 해당 태그를 포함하는 부분-온톨로지를 추출하는 것으로 한다. 따라서 태그 집

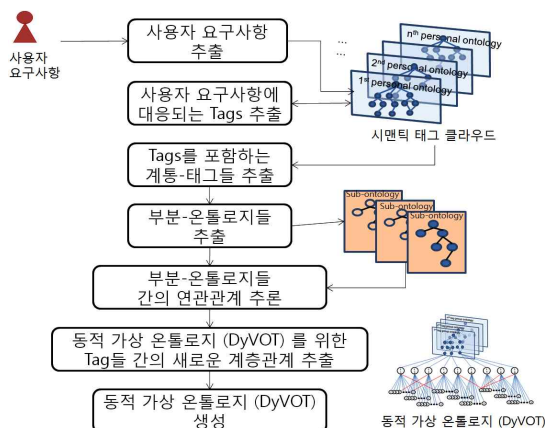
합 T 는 사용자 요구사항 집합 K 로부터 다음과 같이 생성된다.

$$K = \{k_1, k_2, \dots, k_w\}, w \text{은 사용자 요구사항의 수}$$

$$T = \{t_1, t_2, \dots, t_z\}, z \text{은 } K \text{에 대응되는 태그의 수.}$$

태그 그룹은 정적 온톨로지로부터 추출된 태그들 t_i 로 구성되며 태그 그룹을 구성하는 t_i 는 정적온톨로지 상에 존재하는 태그들로 한다.

추출된 태그 집합 T 를 이용해 다음 <Figure 1>과 같이 동적 가상 온톨로지(DyVOT)를 생성한다. DyVOT은 정적온톨로지로부터 태그의 계통분류에 의해 추출된 부분-온톨로지들 간의 공유 자원의 정도에 따라 연관 및 계층관계를 정의하여 생성된다.



<Figure 1> Dynamic Virtual Ontology using Tags (DyVOT) Generation Process

DyVOT의 생성은 사용자의 요구사항에 따라 정적온톨로지로부터 추출된 부분-온톨로지들을 이용하여 가상적으로 이루어진다. 이렇게 생성된 DyVOT을 이용해 사용자가 효율적으로 검색하고자 하는 해를 추출하도록 사용된다. 생성된 동적 가상 온톨로지는 정적온톨로지와의 관계를 형성하여 정적 온톨로지의 진화를 위해 관리되며, 이의 과정에서 변

형 또는 소멸될 수 있다. 이의 생성된 동적 가상 온톨로지의 관리 및 소멸에 대해서는 온톨로지 진화 (Ontology evolution) 연구와 관련하여 다음 연구에서 논의하기로 한다.

다음 장에서는 DyVOT 생성을 위한 부분-온톨로지에 대해 논의한다.

4. 태그를 이용한 부분-온톨로지 생성

부분-온톨로지는 태그 기반 동적 가상 온톨로지 (DyVOT)을 구성하는 기본 단위로 태그 기반 정적 온톨로지로부터 태그 들 중 같은 계통에 있는 태그들을 이용하여 생성된다. 따라서 이를 위해 사용될 사용자 요구사항에 대응되는 태그 집합, 검색 해로서의 자원, 동적 가상 온톨로지 구성을 위한 계통-태그 및 부분-온톨로지 등 각각의 정의 및 생성과정을 설명하고자 한다.

4.1 정적온톨로지와 자원

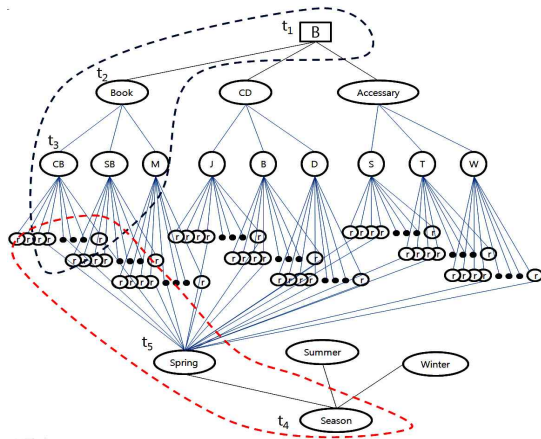
시맨틱 태그 클라우드 기반 정적온톨로지는 클래스와 클래스 간의 상하 관계(Subsumption과 Hyper/ Hypo)로 그 관계가 정의되어 있으며, 온톨로지의 하단인 인스턴스에는 자원이 매칭되어 구성되어 있다(<http://semanticcloud.sandra-siegel.de/>). 또한 태그와 자원 간에는 다대다 대응이 가능하다. 즉 한 태그는 여러 자원에 매칭될 수 있으며, 한 자원 또한 여러 태그에 의해 매칭 될 수 있다. 태그집합을 구성하는 보다 많은 태그들에 의해 특정 자원이 매칭 될 때 그 자원이 사용자의 요구사항을 만족하는 해가 될 확률이 높아진다.

자원은 정형적 또는 비정형적 형태로 구성되며 웹 상에 분포되어 있는 홈페이지, SNS의 단문, 및 이미지, 동영상 등의 형식으로, 즉 정형적 및 비정형적으로 존재한다.

4.2 계통-태그 및 부분-온톨로지

태그 집합을 구성하는 태그들 중 정적온톨로지에서 같은 분류 가지 상에 존재하는 태그들을 계통-태그라 한다. 계통-태그를 포함하여 추출된 트리는 정적 온톨로지의 최상위(root) 노드로부터 계통-태그들을 모두 포함하여 최말단(leaf) 노드들을 포함하여 구성된다. 계통-태그 등에 의해 추출된 트리들을 이용하여 DyVOT를 구성하기 위한 부분-온톨로지들을 구성한다.

부분-온톨로지는 계통-태그를 포함하는 트리에서 계통-태그에 속하면서 트리의 가장 상위에 위치한 태그를 최상위 노드로 하고, 이로부터 트리에 존재하는 계통-태그에 속한 태그들을 따라 인스턴스(자원)가 대응된 말단(leaf) 노드까지로 정의된다. 따라서, 부분-온톨로지의 수는 추출된 계통-태그의 수에 비례한다.



<Figure 2> Extraction of Branches to Construct Sub-ontology

예를 들면, <Figure 2>와 같이 태그 조합 $\langle t_1, t_2, t_3, t_4, t_5 \rangle$ 이 있고 정적 온톨로지로부터 계통-태그 $\langle t_1, t_2, t_3 \rangle$ 와 $\langle t_4, t_5 \rangle$ 이 각각 다른 경로에 존재한다면 다음과 같은 트리들을 추출할 수 있다. 태그 t_1 ,

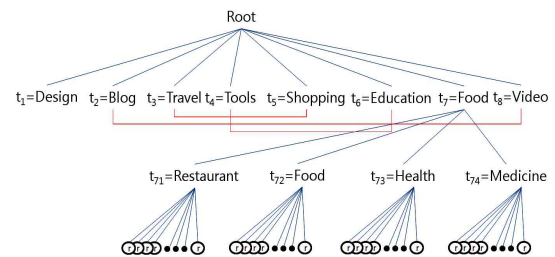
t_2 및 t_3 를 포함하는 트리 $\{\text{root} \rightarrow c_i \rightarrow \dots \rightarrow t_i \rightarrow \dots \rightarrow t_2 \rightarrow t_3 \rightarrow \dots \rightarrow l_{ins_i}\}$ 와 태그 t_4 및 t_5 를 포함하는 트리 $\{\text{root} \rightarrow c_j \rightarrow \dots \rightarrow t_j \rightarrow \dots \rightarrow t_4 \rightarrow \dots \rightarrow t_5 \rightarrow \dots \rightarrow l_{ins_j}\}$ 를 추출할 수 있다. 두 트리에서 각각 t_1 과 t_4 를 최상위 노드로 하는 부분-온톨로지들 O_i 및 O_j 를 추출한다. 추출된 부분-온톨로지들은 인스턴스(l_{ins_i}, l_{ins_j})로서 자원들을 포함한다.

5. 부분-온톨로지들 간의 관계 생성

본 장에서는 부분-온톨로지들 간의 관계 및 관계형성에 대해 논의하기로 한다. 관계는 계층관계와 연관관계로 이루어진다.

정적온톨로지는 태그 클라우드에 포함되어 있는 태그들을 분류 기준에 따라 카테고리화 한다. 같은 카테고리에 속한 태그들 간에는 의미기반 상하 계층(Subsumption과 hyper/hypo) 관계에 의해 정의되며, 이들은 자원들에도 연결될 수 있다.

그러나 본 연구에서 제안하는 동적 가상 온톨로지 DyVOT은 기존의 정적 온톨로지와 달리 <Figure 3>과 같이 태그들 간의 의미적 상하 계층관계뿐만 아니라 태그들 상호 계통(inter-classification) 간에 존재하는 개념적 연관관계까지도 고려한다. <Figure 3>은 계층 및 연관 관계를 고려한 DyVOT을 도시 한다.



<Figure 3> Semantic Hyper/Hypo Relationship and Interrelationship in DyVOT

<Figure 3>에서 의미적 상하 관계로 정의된 계층관계는 'Food'와 'Restaurant'으로 표현된다. 전술한 바와 같이, 연관관계는 태그들 간의 상호 의미적 연관 관계로서, 자원의 공유도에 의해 도출된다. <Figure 3>에서 'Travel'과 'Shopping' 등은 서로 다른 카테고리에 속해 있지만 개념적 연관관계가 전혀 없다고는 할 수 없다. 따라서 두 태그 t_i 와 t_j 간의 연관도인 $AR(t_i, t_j)$ 는 식 (1)과 같이 공유하는 자원의 정도에 따라 계산될 수 있다(Knautz, 2010).

$$AR(t_i, t_j) = 2 \cdot n(t_i, t_j) / (n(t_i) + n(t_j)) \quad (1)$$

t_i, t_j : 태그

$n(t_i)$: 태그 t_i 를 포함하는 자원의 수

$n(t_j)$: 태그 t_j 를 포함하는 자원의 수

$n(t_i, t_j)$: 태그 t_i, t_j 를 모두 포함하는 자원의 수

즉, 두 태그가 공유하는 자원의 수가 많으면 많을수록 연관도가 높다고 본다. 이는 두 태그 간의 개념의 인접도가 높음을 의미한다. 따라서 연관도 값이 임계치(δ) 이상이 되면 두 태그 간에 새로운 연관관계를 생성할 수 있다.

이는 사용자 요구사항에 매칭된 태그들의 종류에 따라 공유하는 자원의 수와 종류가 달라질 수 있음을 의미한다. 같은 태그라도 태그 조합에 따라 태그들 간의 연관도 값이 변할 수 있다. 이를 연관도라 정의하고 이에 따라 태그 간의 동적 연관관계를 정의한다.

6. 동적 가상 온톨로지 DyVOT

DyVOT은 각 태그 조합을 포함하는 계층-태그들에 기반해 생성된 부분-온톨로지들 간의 관계를 정의하여 생성된 동적 가상 온톨로지이다. 즉 사용자의 요구사항에 따라 검색의 편의성과 효율성을 증대하고자 생성된 온톨로지이다.

DyVOT의 생성은 다음과 같은 과정을 거친다. 예를 들어, 계층-태그 $\langle t_1, t_2, t_3 \rangle$, $\langle t_4, t_5 \rangle$ 및 $\langle t_6 \rangle$ 을 포함하는 각각의 부분-온톨로지 O_i, O_j 및 O_k 가 존재한다면, 가장 많은 태그들을 포함하는 계층-태그의 부분-온톨로지 O_i 가 DyVOT을 구성하기 위한 주축 온톨로지가 된다. 주축 온톨로지에 나머지 부분-온톨로지 O_j 및 O_k 를 통합하여 DyVOT을 생성한다.

부분-온톨로지들을 통합하기 위해서는 태그들 간의 연관관계를 계산한다. 즉 태그들 간의 공유하는 자원의 정도가 가장 큰 태그들로 구성된 새로운 클래스를 생성한다. 예를 들어, 주축 온톨로지 O_i 와 나머지 부분-온톨로지 O_j 및 O_k 의 연관관계를 계산하기 위해 주축 온톨로지 상의 태그를 포함하는 가장 하위 클래스들과 부분-온톨로지 O_j 에서 태그를 포함하는 가장 하위 클래스들과의 연관도를 계산한다. 가장 하위 클래스들이 여러 개이면 이들 모두에 대해 각각의 연관도를 계산한다. 주축 온톨로지의 하위 클래스와 O_j 의 하위 클래스 중 연관도가 가장 큰 하위 클래스들 간의 공유하는 자원을 인스턴스로 하여 클래스 ' t_3-t_4 '를 생성한다. 클래스 ' t_3-t_4 '를 t_3 및 t_4 의 하위 클래스로 하는 계층관계를 설정하여, 주축 온톨로지와 부분-온톨로지 O_j 를 통합한다. 즉, 마찬가지로 부분 온톨로지 O_k 의 t_6 는 주축 온톨로지의 ' t_3-t_4 ' 클래스와의 연관관계를 계산하여 연관도가 임계치 이상이면, 새로운 클래스 ' $t_3-t_4-t_6$ '를 생성할 수 있으며, 계층 관계를 정의하여 주축 온톨로지로 통합된다.

만약 주축 온톨로지와 각각의 부분 온톨로지 간의 통합과정에서 이들이 포함하는 태그들 간의 연관도 값이 임계치 이상인 관계가 존재하지 않으면 부분 온톨로지는 독립적 온톨로지로 남는다.

DyVOT 생성과정에서 태그들의 공유자원으로 추론된 자원들과 연관관계가 성립되지 않은 부분 온톨로지에 매칭된 자원들을 해(R)라 하고 식 (2)와 같이 도출한다.

Algorithm for DyVOT

T is a set of tags which are matched into user requirements.

t_i is i^{th} tag.

O_n is n^{th} partial ontology.

b_i is i^{th} branch of Tag.

PO is a primary ontology.

AR is an associative relationship.

HR is a hierarchical relationship.

R is a set of resource as a solution.

DO is a dynamic ontology.

δ is threshold value

// 계통-태그 추출

For all t_i

Set $b_j \leftarrow$ tags are on the same hierarchy path on static ontology

EndFor

// 계통-트리 및 부분-온톨로지 구축

For all b_n

$N_n \leftarrow$ Calculate number of tags of b_n

Set Top-node on $O_n \leftarrow t_i$

// t_i 는 n^{th} b 의 최상위 태그

Set Sub-node on $O_n \leftarrow t_k$

// t_k 는 n^{th} b 의 최하위 태그

For all t_j

Set Middle-node on $O_n \leftarrow t_j$

// t_j 는 n^{th} b 의 j^{th} 태그

EndFor

EndFor

// 주축 온톨로지 추출

Set PO $\leftarrow O_n$ with Max (N_n)

Set S1 $\leftarrow$ A set is composed of the lowest tag of PO

// 계층 및 연관 관계 생성

For all b_i

S2 $\leftarrow$ A set is composed of the lowest tags of sub-node of b_i

For all t_x of S1

For all t_y of S2

Set T1 $\leftarrow t_x$ of S1

Set T2 $\leftarrow t_y$ of S2

Calculate AR_{xy} between T1 and T2

EndFor

EndFor

Set AR $\leftarrow$ Max (AR_{xy})

If AR $> \delta$

then

Create Class ' $t_x - t_y$ '

// 계층관계 생성(t_x 와 ' $t_x - t_y$ ' 및 t_y 와 ' $t_x - t_y$ ')

Set sub-class of $t_x \leftarrow$ class ' $t_x - t_y$ '

Set sub-class of $t_y \leftarrow$ class ' $t_x - t_y$ '

EndFor

// 동적 가상 온톨로지 생성

DO $\leftarrow$ lately constructed ontology with AR and HR relationships

// 해로서의 정보 자원 추출

R $\leftarrow$ resource which is linked to lately constructed class by DO

<Figure 4> DyVOT Generation Algorithm

$$R = \bigcap_{i=1}^n O_i \cup \bigcup_{j=1}^m O_j \quad (2)$$

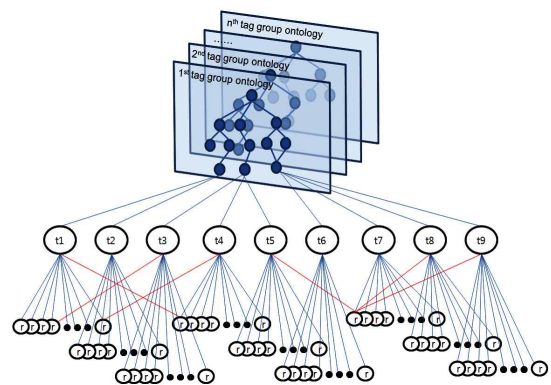
n 은 자원을 공유하는 계통-태그의 수이고 m 은 자원을 공유하지 않는 계통-태그의 수를 나타낸다. 해(R)의 도출은 이들 부분-온톨로지들의 병렬 탐색을 통해 이루어진다. 즉 사용자가 원하는 해를 찾기 위해 각각의 부분-온톨로지들을 이용해 공유하는 자원들과 그렇지 않은 해를 찾아 이를 통합하여 추출한다.

DyVOT을 생성하는 과정은 <Figure 4>의 알고리즘에 기술되어 정의된다.

7. 활용 예제 및 검증

본 논문에서 제안된 DyVOT의 생성과 해의 추출과정을 살펴해보도록 한다.

다음 <Figure 5>는 태그들로 구성된 의미기반 정적온톨로지에 연결된 자원들과 태그들 간의 공유된 자원을 나타낸다. 이들 온톨로지를 이용해 사용자의 요구사항에 따라 탐색을 용이하게 하는 동적 가상 온톨로지 DyVOT을 생성한다.



<Figure 5> Tag-based Static Ontology using Shared Resources

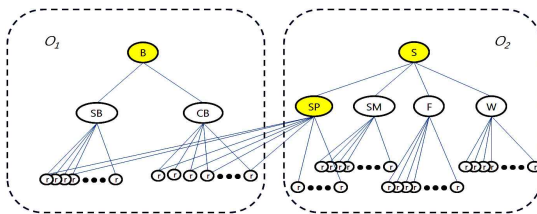
7.1 활용 예제

7.1.1 계통-태그 추출

사용자에 의해 입력된 요구사항이 “봄에 읽고 싶은 책”일 때 이에 대응하는 키워드 집합 K 는 ‘봄,’ ‘책’이 될 수 있다. 또한 이에 매칭된 태그 집합 T 가 {Spring, Season, Book}일 때 정적온톨로지에서의 계층 관계에 있는 태그들끼리 분류하여 계통-태그들 {Book} 및 {Season, Spring}를 추출한다.

7.1.2 계통-트리 및 부분-온톨로지 추출

계통-태그들로부터 트리를 추출한 후, 각 트리에서 사용된 태그들 중 가장 상위에 위치한 태그를 최상위 노드로 하는 부분-온톨로지를 추출한다. 추출된 부분-온톨로지는 $O_1 = \{\text{Season} \rightarrow \text{Spring, Summer, Fall, Winter}\}$ 과 $O_2 = \{\text{Book} \rightarrow \text{Specialty Book, CulturalBook}\}$ 이다.



<Figure 6> Extracted Sub-ontology

<Figure 6>은, ‘B’는 ‘book’, ‘SB’는 ‘specialty books’, ‘CB’는 ‘cultural books,’ ‘S’는 ‘Season’, ‘SP’는 ‘spring’, ‘SM’은 ‘summer,’ ‘F’은 ‘fall’이고 ‘W’은 ‘winter’로 구성된 부분-온톨로지이다.

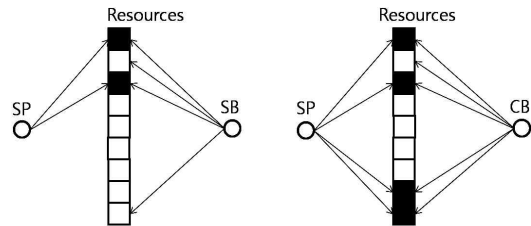
7.1.3 주축 온톨로지 추출

트리 중 태그 집합의 두 개의 태그 ‘Spring,’ ‘Season’를 포함하고 있는 부분-온톨로지가 Dy

VOT 구성의 주축 온톨로지가 된다. 다른 부분-온톨로지는 하나의 태그 ‘Book’만 포함하고 있으므로 더 많은 태그를 포함하고 있는 부분-온톨로지를 주축 온톨로지라 한다.

7.1.4 연관관계 생성

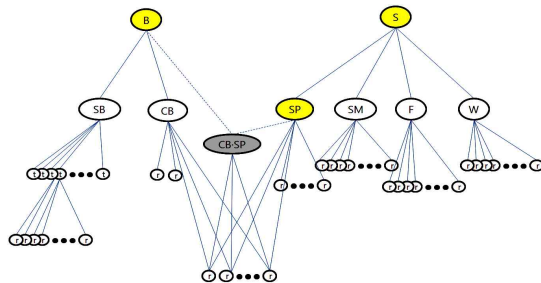
DyVOT 구성을 위해 주축 온톨로지와 나머지 부분-온톨로지와의 태그들 간의 연관관계를 추출한다. 연관관계 계산은 주축 온톨로지의 가장 하위 태그인 ‘SP’와 다른 부분-온톨로지의 태그 ‘B’ 간의 자원 공유 정도에 의해 이루어진다. 태그 ‘B’가 부분 클래스(sub-class)로서 태그 ‘SB’와 ‘CB’를 가지므로 태그 ‘SP’와 ‘SB,’ ‘CB’ 간의 연관관계를 계산한다. <Figure 6>으로부터 ‘SP,’ ‘SB,’ 및 ‘CB’와 자원과 관련된 부분만 추출하면 <Figure 7>과 같다. 각각의 태그에 연관된 자원은 백색 네모로 표현되고 이중 두 태그에 의해 공유된 자원은 흑색 네모로 표현된다. 이때, 식 (1)에 의해 태그 ‘SP’와 ‘CB’간의 자원공유도 0.462가 ‘SP’와 ‘SB’ 간의 자원공유도 0.364보다 크다. 따라서, DyVOT에서 이들 간의 태그 ‘SP’와 ‘CB’로 이루어진 클래스 ‘SP-CB’가 동적으로 생성된다.



<Figure 7> Shared Resources between Tags

7.1.5 계층 관계 생성

태그 ‘SP’와 ‘CB’ 간의 연관관계에 의해 생성된 클래스 ‘SP-CB’는 두 부분-온톨로지들의 태그 ‘SP’와 ‘CB’의 상위 태그들 ‘B’와 관계를 <Figure 8>과 같이 생성한다.



<Figure 8> Generated DyVOT

즉, 태그 ‘SP’는 태그 ‘S’와 의미적관계에 있고, ‘S’와 ‘SP’에 연관된 자원에 연결되므로 하단 클래스 ‘SP’에 연결된 자원은 의미상으로 태그 ‘S’와 ‘SP’에 연관된다. 즉 부분-온톨로지 가장 하단에 있는 태그에 연결된 자원은 계통-태그 상의 태그들과 의미적으로 모두 관련이 있는 자원들이다. 또한 생성된 클래스 ‘CB-SP’는 각각의 부분-온톨로지의 하단의 태그에 연결된다. 즉, O_1 의 ‘B’와 O_2 의 ‘SP’에 연결된다.

7.1.6 동적 가상 온톨로지(DyVOT) 및 해

앞서 일련의 과정을 통해 부분-온톨로지들의 공유자원을 중심으로 생성된 온톨로지를 동적 가상 온톨로지 DyVOT라 하고, 이에 연결된 정보 자원은 사용자의 요구사항에 의해 탐색된 해로 정의된다. 즉, 태그들을 검색의 제약조건으로 하여 검색 범위를 최대한으로 줄여가면서 효율적으로 최적의 자원을 검색한다.

7.2 DyVOT 유용성 검증

검색의 효율을 증명하기 위해, 예제에서 사용된 태그 조합 $T = \{\text{Spring, Season, Book}\}$ 를 이용하여 자원을 검색할 경우 검색할 자원의 개수가 α 개라면 ‘Spring,’ ‘Season’ 및 ‘Book’의 태그를 이용할 경우 검색에 사용되는 비용은 $({}_3C_1+{}_3C_2+{}_3C_3)\times\alpha$ 로 7α 이

상의 검색 비용이 요구된다. 각각의 태그를 포함하는 자원, 2개의 태그를 포함하는 자원 및 3개 태그 모두를 포함하는 자원을 모두 검색한다. 즉, 태그의 수가 n 이라면 태그에 대해 검색에 요구되는 비용은 $({}_nC_1+{}_nC_2+\dots+{}_nC_n)\times\alpha$ 이 요구된다.

따라서 검색의 시간과 비용의 효율을 위해 검색의 제약조건을 상대적으로 강화시켜 검색의 영역을 합리적으로 좁혀 주고 검색의 정확성을 강화할 필요가 있다. 따라서 태그들 중 정적온톨로지 상에서 같은 계통에 존재하는 태그들끼리 분류하여 이들과 연관된 자원을 검색하고, 계통-태그 간의 공유자원을 검색하여 자원의 검색 속도와 정확성을 강화하고자 하였다.

검색에 사용될 태그들을 이용한 태그조합의 경우 <Figure 4>에서 ‘S’(Season)의 ‘SP’(Spring)은 같은 태그 계통으로 계층관계에 있으므로 이를 (S, SP)로 하는 하나의 태그 그룹으로 할 수 있다. 따라서 검색해야 할 태그 집합은 실제로 {B, (S, SP)}로 볼 수 있고, 탐색해야 할 태그의 수는 $n = 2$ 이므로 $n = 3$ 일때의 검색 비용 $({}_3C_1+{}_3C_2+{}_3C_3)\times\alpha = 7\alpha$ 보다 적은 $({}_2C_1+{}_2C_2)\times\alpha = 3\alpha$ 의 검색 비용이 요구된다. 즉, 4α 이상의 검색 비용을 줄일 수 있다. 예를 들어, 자원의 수에 해당하는 값이 10,000개일 경우 40,000개 이상의 자원에 대한 검색 비용이 줄어 들게 된다. <Figure 8>의 경우 ‘B’의 ‘CB’와 (S, SP)간에 새로운 클래스 ‘CB-SP’를 형성하므로 탐색해야 할 태그의 수 $n=1$ 이 되어 6α 이상의 비용을 줄일 수 있다. 즉, 태그를 이용한 동적 가상 온톨로지 DyVOT을 이용해 일반적 키워드 검색에 비해 검색 대상의 양을 줄여 검색의 비용을 줄임과 동시에 검색 정확성을 높일 수 있다.

8. 결론

본 연구에서는 대용량의 정보나 문서로부터 사용

자 요구사항에 부합하는 가치 있는 정보의 효과적 추출을 위해서 태그들 간의 의미적 관계를 이용하여 검색 대상을 합리적으로 줄이는 동적 가상 온톨로지를 이용한 정보 검색 방법론을 제안하고자 하였다. 이를 위해 태그를 이용하여 태그들 간의 의미적 계층 및 연관관계를 정의한 동적 가상 온톨로지 DyVOT을 작성하여 검색의 효율을 증대시키고자 하였다. 즉 사용자의 요구사항에 매칭되는 태그들을 이용하여 사용자가 필요로 하는 정보자원을 효율적 검색을 지원하고자 제안되었다. 데이터의 의미론적 분류를 위해 태그를 이용한 것은 데이터 생성 및 사용에 사용자들이 적극적인 참여가 가능해졌고, 태그 또한 사용자에게 의해 정의된다는 점에 있어 사용자 참여에 기반한 데이터 분류는 의미 있다고 보았다. 또한 기존의 시맨틱 클라우드가 단순히 태그 클라우드에 의미론에 기반한 데이터 분류에 초점을 두었다면 본 연구에서는 의미론적으로 분류된 태그들 간의 연관관계를 정의하여 온톨로지를 구축하였다. 또한 구축된 DyVOT은 태그의 의미 기반 온톨로지를 이용한 클러스터링에 기반해 검색 대상을 줄여 기존의 키워드 기반 검색에 비해 검색 비용을 현저히 줄여 주었으며 데이터의 정확도를 증가시켜 주는 효과가 있었다. 다음 연구에서는 생성된 동적 가상 온톨로지를 기존의 정적온톨로지에 반영하여 온톨로지의 진화(evolution) 및 확장에 대한 논의를 계속하여 소개하고자 한다.

참고문헌

- Angeletou, S., M. Sabou, L. Specia, and E. Motta, "Bridging the gap between folksonomies and the semantic web : An experience report", *Workshop: Bridging the Gap between Semantic Web and Web 2.0, European Semantic Web Conference*, (2007), 30~43.
- Arias, J. A. and J. A. Sánchez, "Content-Based Search and Annotations in Multimedia Digital Libraries", *Proceeding ENC '03 Proceedings of the 4th Mexican International Conference on Computer Science, IEEE Computer Society* (2003).
- Carter, P., "Big Data" to Drive the Next Wave of Investments in Business Analytics for CIOs in Asia/Pacific, IDC Analyze the Future, 2011.
- Dean, J. and S. Ghemawat, "MapReduce : Simplified Data Processing on Large Clusters", *OSDI*(2004).
- Dean, J. and S. Ghemawat, "MapReduce: Simplified Data Processing, on Large Clusters", *COMMUNICATIONS OF THE ACM*, Vol. 51, No.1(2008).
- Ganguly, S., S. Bhatnagar, A. Saxena, S. Banerjee, and R. Izmailov, "A Fast Content-based Data Distribution Infrastructure", *INFOCOM 2006 25th IEEE International Conference on Computer Communications*, Proceedings(2006).
- Halvey, M. and M. Keane, "An assessment of tag presentation techniques", *Proceedings of the 16th International Conference on World Wide Web*, (2007), 1313~1314.
- Hassan-Monteroa, Y. and V. Herrero-Solanaa, "Improving Tag-Clouds as Visual Information Retrieval Interfaces", *International Conference on Multidisciplinary Information Sciences and Technologies*, InSciT2006. Mérida, Spain. October(2006), 25~28.
- Hotho, A., R. Jäschke, C. Schmitz, and G. Stumme, "Information Retrieval in Folksonomies : Search and Ranking", *The Semantic Web : Research and Applications*, Vol.4011 of LNAI, (2006).
- Jäschke, R., L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme, "Tag Recommendations in Folksonomies", *Web Tagging Workshop at the WWW*(2006).

- Jung, J. J., "Collaborative browsing system based on semantic mashup with open APIs", *Expert Systems with Applications*, Vol.39(2012), 6897~6902.
- Jung, J. J., "Discovering Community of Lingual Practive for Matching Multilingual Tags from Folksonomies", *The Computer Journal*, (2011).
- Knautz, K., S. Soubusta, and W. G. Stock, "Tag Clusters as Information Retrieval Interfaces", *Proceedings of the 43rd Hawaii International Conference on System Sciences*, (2010).
- Lv, Q., W. Josephson, Z. Wang, M. Charikar, and K. Li, "Ferret: A Toolkit for Content Based Similarity Search of Feature Rich Data", *Euro Sys '06*, April, (2006), 18~21.
- Madden S., M. J. Franklin, J. Hellerstein, and W. Hong, "TAG : a Tiny Aggregation Service for Ad-Hoc Sensor Networks", *Proceedings of the Fifth Symposium on Operating Systems Design and implementation(OSDI '02)*, (2002).
- Markines, B., C. Cattuto, F. Menczer, D. Benz, A. Hotho, and G. Stumme, "Evaluating similarity measures for emergent semantics of social tagging", *Proceedings of the 18th International Conference on World Wide Web*, (2009).
- McKinsey Global Institute, *Big data: The next frontier for innovation, competition, and Productivity*, MacKinsey and Company, 2011.
- Park, J., N. Kim, M. Choi, Z. Jin, and Y. Choi, "Semantic Search : A Survey", *Journal of Intelligence and Information Systems*, Vol.17, No.4(2011), 19~36.
- Riabov, A. V., E. Bouillet, M. D. Feblowitz, Z. Liu, and A. Ranganathan, "Wishful Search : Interactive Composition of Data Mashups", *WWW 2008/Refereed Track: Web Engineering-Web Service Composition, WWW 2008*, April, (2008), 21~25.
- Specia, L. and E. Motta, "Integrating Folksonomies with the Semantic Web", *Lecture Notes in Computer Science*, Vol.4519(2007), 624~639.
- Salonen, J., "Self-organising map based tag clouds: Creating spatially meaningful representations of tagging data", *OPAALS conference*, (2007).
- Schrammel, J., M. Leitner, and M. Tscheligi, "Semantically structured tag clouds: An empirical evaluation of clustered presentation approaches", *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, (2009), 2037~2040.
- Yang, B. and A. R. Huron, "*Semantic-aware data processing: towards cross-modal multimedia analysis and content-based retrieval in distributed and mobile environments*", *Ph.D. Dissertation*, Dept. of Computer Science and Engineering, Pennsylvania State University, (2007).
- Yoo, K., "Ontology-Based Process-Oriented Knowledge Map Enabling Referential Navigation between Knowledge", *Journal of Intelligence and Information Systems*, Vol.18, No.2(2012), 61~83.
- Yoo, D., G. Kim, K. Choi, and Y. Suh, "CTKOS: Categorized Tag-based Knowledge Organization System", *Journal of Intelligence and Information Systems*, Vol.17, No.4(2011) 59~74.
- Zhang, L., X. Wu, and Y. Yu, "Emergent semantics from folksonomies: A quantitative study", *Journal on Data Semantics VI(Lecture Notes in Computer Science)*, (2006), 168~186.
- Zikopoulos, P. C., C. Eaton, D. Deroos, T. Deutsch, G. Lapis, *Understanding big data Analytics for Enterprise Class Hadoop and Streaming data*, McGraw Hill, (2012).
- Carter, P., "*Big Data*", *to Drive the Next Wave of Investments in Business Analytics for CIOs in Asia/Pacific*, IDC Analyze the Future, (2011).

Abstract

Dynamic Virtual Ontology using Tags with Semantic Relationship on Social-web to Support Effective Search*

Hyun Jung Lee** · Mye Sohn***

In this research, a proposed Dynamic Virtual Ontology using Tags (DyVOT) supports dynamic search of resources depending on user's requirements using tags from social web driven resources.

It is general that the tags are defined by annotations of a series of described words by social users who usually tags social information resources such as web-page, images, u-tube, videos, etc. Therefore, tags are characterized and mirrored by information resources. Therefore, it is possible for tags as meta-data to match into some resources. Consequently, we can extract semantic relationships between tags owing to the dependency of relationships between tags as representatives of resources.

However, to do this, there is limitation because there are allophonic synonym and homonym among tags that are usually marked by a series of words. Thus, research related to folksonomies using tags have been applied to classification of words by semantic-based allophonic synonym. In addition, some research are focusing on clustering and/or classification of resources by semantic-based relationships among tags. In spite of, there also is limitation of these research because these are focusing on semantic-based hyper/hypo relationships or clustering among tags without consideration of conceptual associative relationships between classified or clustered groups. It makes difficulty to effective searching resources depending on user requirements.

In this research, the proposed DyVOT uses tags and constructs ontology for effective search. We assumed that tags are extracted from user requirements, which are used to construct multi sub-ontology

* This research was partially supported by the MKE(The Ministry of Knowledge Economy), Korea, under the "IT Consilience Creative Program" support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2012-H0201-12-1001) and partially supported by Defense Acquisition Program Administration and Agency for Defense Development under the contract UD080042AD, Korea.

** Yonsei Institute of Convergence Technology, School of Integrated Technology, Yonsei University

*** Corresponding Author: Mye Sohn

Department of Industrial Engineering, Sungkyunkwan University

300 Cheoncheon-dong, Jangan-Gu, Suwon, Korea

Tel: +82-2-290-7605, Fax: +82-2-290-7610, E-mail: myesohn@skku.edu

as combinations of tags that are composed of a part of the tags or all. In addition, the proposed DyVOT constructs ontology which is based on hierarchical and associative relationships among tags for effective search of a solution. The ontology is composed of static- and dynamic-ontology. The static-ontology defines semantic-based hierarchical hyper/hypo relationships among tags as in (<http://semanticcloud.sandra-siegel.de/>) with a tree structure. From the static-ontology, the DyVOT extracts multi sub-ontology using multi sub-tag which are constructed by parts of tags. Finally, sub-ontology are constructed by hierarchy paths which contain the sub-tag.

To create dynamic-ontology by the proposed DyVOT, it is necessary to define associative relationships among multi sub-ontology that are extracted from hierarchical relationships of static-ontology. The associative relationship is defined by shared resources between tags which are linked by multi sub-ontology. The association is measured by the degree of shared resources that are allocated into the tags of sub-ontology. If the value of association is larger than threshold value, then associative relationship among tags is newly created. The associative relationships are used to merge and construct new hierarchy the multi sub-ontology. To construct dynamic-ontology, it is essential to defined new class which is linked by two more sub-ontology, which is generated by merged tags which are highly associative by proving using shared resources. Thereby, the class is applied to generate new hierarchy with extracted multi sub-ontology to create a dynamic-ontology. The new class is settle down on the ontology. So, the newly created class needs to be belong to the dynamic-ontology. So, the class used to new hyper/hypo hierarchy relationship between the class and tags which are linked to multi sub-ontology. At last, DyVOT is developed by newly defined associative relationships which are extracted from hierarchical relationships among tags. Resources are matched into the DyVOT which narrows down search boundary and shrinks the search paths. Finally, we can create the DyVOT using the newly defined associative relationships.

While static data catalog (Dean and Ghemawat, 2004; 2008) statically searches resources depending on user requirements, the proposed DyVOT dynamically searches resources using multi sub-ontology by parallel processing. In this light, the DyVOT supports improvement of correctness and agility of search and decreasing of search effort by reduction of search path..

Key Words : Tag, Search, Dynamic Virtual Ontology, Semantics Social-web Resource

저 자 소 개



이현정

이화여자대학교 물리학사 및 전자계산학과 석사를 받았으며, 한국과학기술원 (KAIST) 테크노경영대학원 경영공학 박사학위를 취득하였다. 현재 연세대학교 미래융합기술연구원에 재직 중이다. 주요 연구분야는 지능정보시스템, 의사결정지원시스템, 유비쿼터스 컴퓨팅, 시맨틱 웹서비스, 헬스서비스 시스템 및 소셜 웹기반추천시스템 등이다.



손미애

성균관대학교 산업공학과를 졸업하고 한국과학기술원에서 경영정보공학으로 박사학위를 취득한 후, 한국국방연구원에 근무했으며, 현재 성균관대학교 시스템경영공학과 부교수로 재직 중이다. 주요 관심분야는 시맨틱웹, 시맨틱 웹서비스, BPMS 및 Enterprise Architecture/Architecture Framework 등이며, 현재 온톨로지상에서의 유사도계산과 SBA를 위한 EA 구축연구를 수행하고 있다. 한국경영공학회와 ITAEA 학회의 임원으로 활동 중이며, CACM을 비롯한 국내외 우수저널에 논문을 발표하고 있다.