

가상 커뮤니티에서 사회 관계 추론을 위한 시맨틱 웹 접근 방법

이승훈^a, 김지혁^a, 김홍남^a, 조근식^b

^a 인하대학교 정보공학과

인천 남구 용현동 253, 402-751

Tel: +82-32-875-5863, Fax: +82-32-875-5863, E-mail: {shlee, jk13, nami}@eslab.inha.ac.kr

^b 인하대학교 컴퓨터 정보공학부

인천 남구 용현동 253, 402-751

Tel: +82-32-860-7440, E-mail: gsjo@inha.ac.kr

요약

최근 Web 2.0 발달과 더불어 블로그나 온라인 카페 등 웹 상의 사용자가 개인적인 정보를 자유롭게 게재할 수 있도록 하는 인터넷서비스가 증가하면서, 이 사용자들 간의 관계에 초점을 맞춘 소셜 네트워크 분야의 연구가 활발히 이루어지고 있다. 하지만 많은 소셜 네트워크 서비스가 정보자원을 컴퓨터가 처리할 수 있는 의미적인 정보로 표현하고 있지 않기 때문에 서로 다른 도메인 간에 공유와 재사용이 어렵고, 사회적 개체들 간의 관계가 명확하게 정의되어 있지 않아 소셜 네트워크 분석에 어려움이 있다.

본 논문에서는 가상 커뮤니티 사용자들이 업로드한 사진 데이터를 이용한 시맨틱 웹 기반의 소셜 네트워크 분석 시스템을 제안한다. 온톨로지를 기반으로 사진에서 추출된 얼굴 개체와의 관계와 이미 인맥 관계를 형성하고 있는 사람들의 정보적 연결성을 명확하게 정의하고 도메인 규칙을 활용하여 의미 있는 사회적 연결 관계를 추론한다. 이를 그래프로 시각화하여 사용자에게 제공함으로써 온라인 상에서 형성된 커뮤니티 내에서 효율적인 소셜 네트워크 분석을 도모하고 이를 기반으로 다양한 응용 분야에 활용하는 방법을 모색한다.

키워드:

가상 커뮤니티; 시맨틱 소셜 네트워크 분석; 소셜 온톨로지;

1. 서론

최근 참여와 개방을 핵심으로 한 Web 2.0과 더불어 웹 상의 정보 공간에 사용자가 자신의 개인적인 정보를 자유롭게 게재할 수 있도록 하는 소셜

네트워크 서비스가 증가하였다. 그리고 MySpace¹, facebook², 싸이월드³ 등 인기 있는 소셜 네트워크 사이트는 엄청난 규모의 트래픽과 커뮤니티 회원을 확보하고 있다. 따라서 이러한 가상의 커뮤니티 사용자들 간의 관계에 초점을 맞춘 소셜 네트워크 분야의 연구가 활발히 진행되고 있고, 소셜 네트워크 분석을 통하여 새로운 부가가치 창출을 하기 위한 많은 노력이 이루어지고 있다.

소셜 네트워크 분석(Social Network Analysis)은 주어진 범위 내의 전체 관계에서 사용자와 이미 인맥관계를 형성하고 있는 사람들의 정보적 연결성을 통해 새로운 사회적 연결 관계를 발견해 내는 것에 목적이 있다[4, 5]. 그러나 현재 서비스되고 있는 대부분의 소셜 네트워크 사이트는 사용자와 관련된 정보자원 및 개체들 간의 관계에 대하여 명확하게 정의하고 있지 않아 의미적 관계를 발견하지 못하는 한계가 있다. 이러한 문제점은 사회적 개체와 이들 간의 관계에 관련된 도메인 온톨로지(ontology), 적절한 규칙(rule), 그리고 추론 메카니즘을 통해 새로운 정보를 추론하여 새로운 관계나 이전에 존재하지 않았던 새로운 개념을 발견해 냄으로써 극복 될 수 있다[2].

본 논문에서는 이러한 기존의 소셜 네트워크 서비스의 단점을 보완하기 위하여, 시맨틱 웹 기반의 소셜 네트워크 분석 시스템을 제안한다. 사용자가 업로드한 개인의 사진 분석을 통하여 의미적인 가상 커뮤니티를 반자동적으로 구성하고 시각화하며, 추론을 통하여 새로운 의미적 관계를 발견해 낸다. 이를 위해 제안된 시스템의 디자인 목표는 다음과 같다. 첫 번째, 의미적으로 잘 정의된 개인 사진 관리를 제공한다. 두 번째, 사용자들의

¹ <http://www.myspace.com/>

² www.facebook.com/

³ <http://www.cyworld.com/>

관계 및 사용자와 사진과의 명확한 관계를 정의한다. 세 번째, 의미적인 가상 커뮤니티(소셜 네트워크)를 분석한다. 마지막으로, 정보자원(resource)의 재사용과 공유를 증진시킨다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구로 기존의 연구들을 설명하고, 3장에서는 본 논문에서 제안하는 얼굴 사진을 이용한 소셜 네트워크 분석 시스템에 대해 자세히 기술한다. 4장과 5장에서는 각각 구현된 프로토타입과 실험결과에 대해서 분석, 평가하고, 마지막으로 6장에서 결론을 맺고 향후 연구에 대하여 언급한다.

2. 관련연구

2.1. 온톨로지를 기반으로 한 소셜 네트워크 분석

소셜 네트워크란 일반적으로 친구, 직장 동료 또는 정보를 교환하는 관계 등 다양한 사회적 관계를 통해 연결된 사람이나 조직 혹은 다른 사회적 개체의 집합을 의미한다[1]. 최근 인터넷의 발달과 더불어 블로그나 온라인 카페 등 웹 상의 네트워크화 된 정보 공간에 사용자가 자신의 개인적인 정보를 자유롭게 게재할 수 있도록 하는 서비스가 증가하면서, 이 사용자들 간의 관계에 초점을 맞춘 소셜 네트워크 분야의 연구가 활발히 이루어지고 있다. 이와 같은 사용자들은 단순히 사회적인 측면뿐만 아니라 교육, 정치, 경제 등 다양한 분야의 가상 커뮤니티를 형성함으로써 현대 사회의 주요한 부분으로 자리매김하고 있다[3].

소셜 네트워크 분석(Social Network Analysis: SNA)은 이러한 사회적 개체들 간의 관계를 바탕으로 이미 알려져 있지 않은 특정한 패턴을 발견해내는 데 그 목적이 있다[4, 5]. 이미 언급한 바와 같이 폭넓은 분야에서 연구가 계속되고 있으며, 이에 따른 다양한 응용 연구가 진행되어 왔다. C. Chen은 전자 도서관 자료의 저자들이 참조한 타 자료의 목록을(co-citation) 바탕으로 이들의 관계를 분석하고 시각화하였다[6]. 또한 문서의 공동 저자간의 관계를 분석하고[7], 더 나아가 공동 저자 네트워크 내에서 영향력 있는 개인을 발견해 내기도 한다[8]. 상용되고 있는 인터넷 소셜 네트워킹 서비스로는 주로 전문가 집단의 네트워크를 형성하기 위한 비즈니스 지향적 소셜 네트워킹 사이트인 LinkedIn⁴, Linknow⁵ 등이 있으며, orkut⁶, MySpace, facebook, 싸이월드 등과 같이 개인적인 친분 관계의 유지와 확장, 정보의 공유를 목적으로 소셜 네트워크를 구성하는 서비스 등이 있다.

연구와 적용 분야의 다양성만큼이나 분석을 위한

접근 방법 또한 매우 다양한데, 최근 시멘틱 웹 패러다임의 발달과 함께 소셜 네트워크 모델링과 분석의 접근 방법 중 하나로 온톨로지의 이용이 큰 비중을 얻고 있다.

P.O.Wennerberg는 소셜 네트워크의 모델링과 분석에 온톨로지를 사용함으로써 얻을 수 있는 이점을 다음과 같이 설명하였다[2]. 우선 온톨로지는 일반적으로 주어진 도메인에 관련된 개념과 이들 상호간의 관계를 설명하고 명세를 제공하기 위해 사용된다. 소셜 네트워크는 각 개체와 개체간의 사회적 관계에 대한 명세를 제공하고자 하는 목적을 갖는다는 점에서 온톨로지의 일반적인 목적과 부합한다. 또한, 온톨로지는 양립되거나 일관성 없는 정보의 모델링을 허용하지 않기 때문에 정보의 유효성을 보장할 수 있다. 마지막으로, 온톨로지는 적절한 규칙(rule)과 추론 메카니즘을 통해 새로운 정보를 추론하여 얻을 수 있도록 한다. 이전에 존재하지 않았던 개체 간의 새로운 관계를 발견해 내는 데 이러한 추론 메카니즘을 사용할 수 있다.

온톨로지를 이용한 소셜 네트워크의 분석과 이에 관련한 응용 사례는 다음과 같다. P. Mika는 온톨로지로 표현되는 소셜 네트워크 데이터를 수집하고 이를 이용하여 소셜 네트워크를 분석, 평가하였으며[9], P.O.Wennerberg는 온톨로지 및 소셜 네트워크의 분석을 통해 특정 도메인의 지식 베이스에서 지식을 발견(knowledge discovery)해내는 하나의 접근방법으로 사용하였다[2]. 또한 [3]에서는 두 개의 서로 다른 소셜 네트워크를 통합하는 과정에서 발생하는 이익 충돌(conflict of interest) 관계를 효과적으로 분석하기 위해 온톨로지를 사용하였고, 사용자에게 적절한 정보를 추천해 주기 위한 목적으로 온톨로지를 이용하여 사용자를 모델링하고 이들 간의 관계를 분석하는 등의 연구가 진행 중이다[11].

2.2. 사진을 이용한 소셜 네트워크 분석

인터넷 상에서 사진의 효과적인 관리와 공유를 돕는 서비스나 사진 정보를 손쉽게 게시할 수 있는 개인 블로그 형태의 웹 페이지를 제공하는 서비스를 접하기란 그리 어려운 일이 아니다. 대표적인 서비스인 Flickr⁷는 사진을 공유하고자 하는 다른 사용자를 사진의 소유자가 선택할 수 있어, 사진 교환이나 공유 관계를 통해 단순한 형태의 소셜 네트워크 분석이 가능하다.

사진 데이터를 소셜 네트워크 분석에 사용할 경우, 사진에 나타난 사람 간의 관계와 사진의 소유 관계 등에서 소셜 네트워크를 유추할 수 있으며, 사진에 대한 메타 데이터를 주석(annotation)이나 태그 등을 붙이는 방법으로 다양한 네트워크의 분석에 사용될 수 있는 풍부한 개체의 정보를 제공하기도 한다.

⁴ <http://www.linkedin.com/>

⁵ <http://www.linknow.kr/>

⁶ <http://www.orkut.com/>

⁷ <http://www.flickr.com/>

riya⁸, fotolog⁹, olalog¹⁰ 등은 사진 속의 특정 개체 영역을 추출하여 사진 데이터의 분류나 검색이 용이하도록 돕는 서비스이다. 특히 olalog는 얼굴 인식 엔진을 바탕으로 개인의 사진으로부터 인물 정보를 추출하여 자동으로 분류하고 공유할 수 있게 해주는 소셜 네트워크 서비스를 제공한다.

이와 같이 사진 데이터를 이용하여 사진 속의 개체나 소유자들 간의 사회적 관계를 분석하여 소셜 네트워크를 구축하고, 이를 사진의 검색이나 분류를 더욱 효율적으로 수행하기 위한 접근 방법의 하나로 활용하는 등의 연구가 활발히 진행되고 있다[16, 17, 18, 19, 20, 21]. FamiliarFace[16]는 사용자가 직접 선택한 각 사진의 인물에 대한 관계를 나타내는 주석을 바탕으로, 사진으로 구성된 가족 관계 트리를 자동으로 구성해주는 프로토타입 어플리케이션이다. 본 논문에서 제안한 시스템 구조와 매우 유사하나 생성된 소규모의 소셜 네트워크를 바탕으로 더욱 다양한 관계의 추론과 분석 등이 충분히 기술되고 있지 않다. [19]는 개인적인 정보와 상황, 사회적 정보와 상황 등에 기초하여 사진에 주석을 첨부하기 위한 프레임 워크로, 소셜 네트워크에서 발생할 수 있는 사회적 사건의 상호관계를 수량적으로 측정한다. 이 밖에도 Flickr의 소셜 네트워크를 확장하여 사회적 정보를 분석하고 이를 바탕으로 사용자에게 사진을 추천하거나[20], 사진의 메타 데이터로서 얻어지는 소셜 네트워크 정보를 분석하여 사진의 효율적인 분류와 검색에 활용하는 예 또한 다양하다[17, 18].

2.3. Friend Of A Friend

FOAF 프로젝트는 사람에 관련된 정보와 그들의 관계, 흥미, 활동 등에 관한 메타 데이터를 표현하기 위해 RDF 어휘(vocabulary)를 정의하는 커뮤니티 기반 연구 활동이다[12].

일반적인 상용 서비스에서 소셜 네트워크를 구성하기 위해 중앙 서버에 이를 사용자의 정보를 저장하고 유지하는 것과는 달리 FOAF 프로파일은 개개의 사용자에게 의해 작성되고, 사용자 개개의 웹 페이지 등에 분산되어 저장된다. 이와 같은 FOAF 프로파일은 또한 각 개인의 프로파일 내에서 자신과 관계있는 사람의 프로파일을 연결(link) 하여 네트워크를 구조화 할 수 있도록 하고, 그 자체로 기계가 처리할 수 있는 공통적인 형식으로 기술된 하나의 소셜 네트워크가 형성 된다[9].

FOAF 문서는 개체의 구별과 기술이 간단하며, 서로 다른 문서간의 통합이 쉽다는 특성을 갖기 때문에, 최근 온라인 소셜 네트워크 분석의 데이터 기술과 분석 방법으로 많이 사용되고 있다. L. Ding et

al.은 FOAF 문서 자체의 구조와 특성을 분석하기 위해, 실험적 이용을 통해 자주 사용되는 어휘, 도출된 소셜 네트워크에서 발견된 흥미로운 연결 패턴 등을 분석하였다[14]. 이를 바탕으로 T. Finin et al. 은 소셜 네트워크를 이용하여 지능형 웹 기반 정보 시스템에 활용할 수 있는 가능성에 관하여 논하였다[10]. 단순한 소셜 네트워크 분석에서 더 나아가 [3]에서는 FOAF로 구성된 소셜 네트워크와 DBLP¹¹ 목록에서 제공하는 저자 간의 공동 저자 관계로 이루어진 협업 네트워크를 결합하고, 이로 인해 발생하는 두 개의 소셜 네트워크 간의 이익 충돌을 탐지하였다. 방대한 소셜 네트워크에서 여러 형태의 사회 그룹을 분석해내기 위해 FOAF로 이루어진 데이터의 기계 학습 방법이 제안되기도 하였다[16, 17]. 또한, 사용자들 간의 신뢰도 정도를 소셜 네트워크의 분석에 활용하기 위해, J. Golbeck은 신뢰 네트워크 모델링을 위한 새로운 FOAF 클래스와 9 단계의 신뢰 정도를 나타내는 Person 클래스의 새로운 속성을 정의하였다[15].

3. 얼굴 사진을 이용한 소셜 네트워크

소셜 네트워크 분석(Social Network Analysis)의 목적은 특정한 사람, 그룹 또는 조직의 관계를 분석하는 것이다[4, 5]. 소셜 네트워크 분석은 주어진 범주 내의 전체 관계에서 사용자와 이미 인맥관계를 형성하고 있는 사람들의 정보적 연결성을 통해 새로운 사회적 연결 관계를 발견해 내는 것이다. 이러한 관계를 분석하기 위해서는 기본적으로 특정 개인이 가지는 일반적인 정보를 수집함으로써 상황이나 범위를 구체화할 필요가 있다. 이러한 정보를 기반으로 사람들 간의 연쇄적 연결 구조를 표현하여 광범위한 소셜 네트워크를 형성할 수 있다. 그러나 단지 사람 사이의 연결 관계를 표현할 수 있을 뿐 관계가 갖는 의미를 발견할 수 있는 것은 아니다.

시맨틱 웹과 소셜 네트워크의 합성어인 시맨틱 소셜 네트워크는 이러한 사회적 연결 관계에 의미를 부여한다. 리소스와 리소스에 관련된 사람이 함께 연결되어 있기 때문에, 사람 간의 다양한 의미적 연결 관계의 표현과 추론이 가능한 공간이라 할 수 있다.

본 연구에서는 사용자들의 사진 분석을 이용하여 시맨틱 소셜 네트워크를 구성하고 새로운 소셜 관계를 추론한다.

⁸ <http://www.riya.com/>

⁹ <http://www.fotolog.com/>

¹⁰ <http://www.olalog.com/>

¹¹ Digital Bibliography & Library Project <http://dblp.uni-trier.de/>

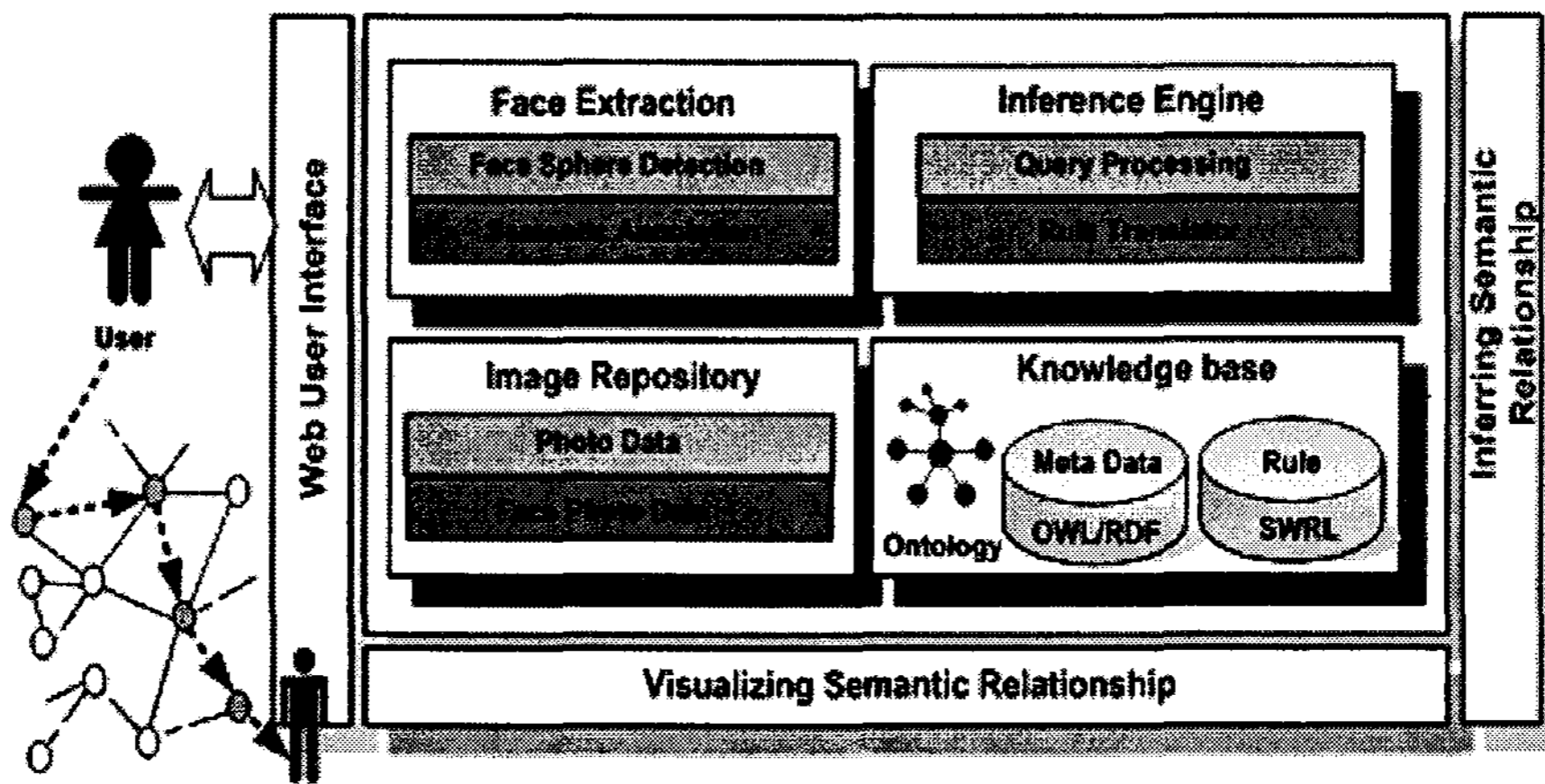


그림 1. 시맨틱 소셜 네트워크 시스템 구조도

3.1. 시맨틱 소셜 네트워크 시스템 구조도

본 논문에서 제안한 얼굴 사진을 이용한 시맨틱 소셜 네트워크 시스템의 구조는 그림 1과 같고, 크게 네 부분으로 나누어 설명할 수 있다.

얼굴 추출 모듈은 그 첫 번째 과정으로, 웹 인터페이스를 통해 사용자가 사진을 업로드하면 얼굴 영역 인식 모듈은 그 사진 속에 등장하는 인물들의 얼굴 영역을 인식하여 추출한다. 사용자가 모듈에 의해 추출된 얼굴 이미지들을 개별 사용자들과 연결함으로써 얼굴 이미지는 그 식별이 가능한 고유성을 가지게 된다.

이미지 저장소에는 사용자가 시스템에 업로드 한 사진과 사진에서 추출된 얼굴 이미지가 저장된다.

지식 베이스 모듈에는 온톨로지와 도메인 규칙으로 표현된 지식과 메타데이터가 위치한다. 온톨로지는 사람, 사진, 사진에서 추출된 얼굴 이미지, 소셜 관계 등 소셜 네트워크를 구축하는데 필요한 정보들을 클래스와 속성으로 정의하여

OWL로 표현된다. 그리고 정의된 클래스와 속성을 바탕으로 SWRL(Semantic Web Rule Language)를 이용하여 도메인 규칙을 생성한다[23].

온톨로지와 얼굴 추출 모듈은 다음과 같은 과정을 통하여 상호 운용된다. 얼굴 추출 모듈을 통해 식별이 가능한 메타데이터로 표현된 얼굴 이미지는 온톨로지의 얼굴 관련 클래스 내에 인스턴스로 포함되고, 사람에 관해 묘사한 클래스와 정의된 속성에 의한 관계를 통해 어떤 개인을 묘사하고 있는지를 발견한다.

마지막으로 추론 엔진은 사용자가 던진 질의를 분석, 처리하고 지식베이스의 인스턴스와 정의된 규칙을 기반으로 새로운 소셜 관계를 추론한다. 그리고 그 결과를 사용자에게 시각화하여 제공한다.

3.2. 소셜 네트워크 분석을 위한 온톨로지

사용자들에 의해 업로드 된 사진에 관한 정보, 개인이 가지는 프로필, 사람 간의 연결 관계에 대한 소셜 온톨로지를 OWL로 정의한다. 그림 2는 온톨로지의 클래스와 속성의 개략적인 관계를

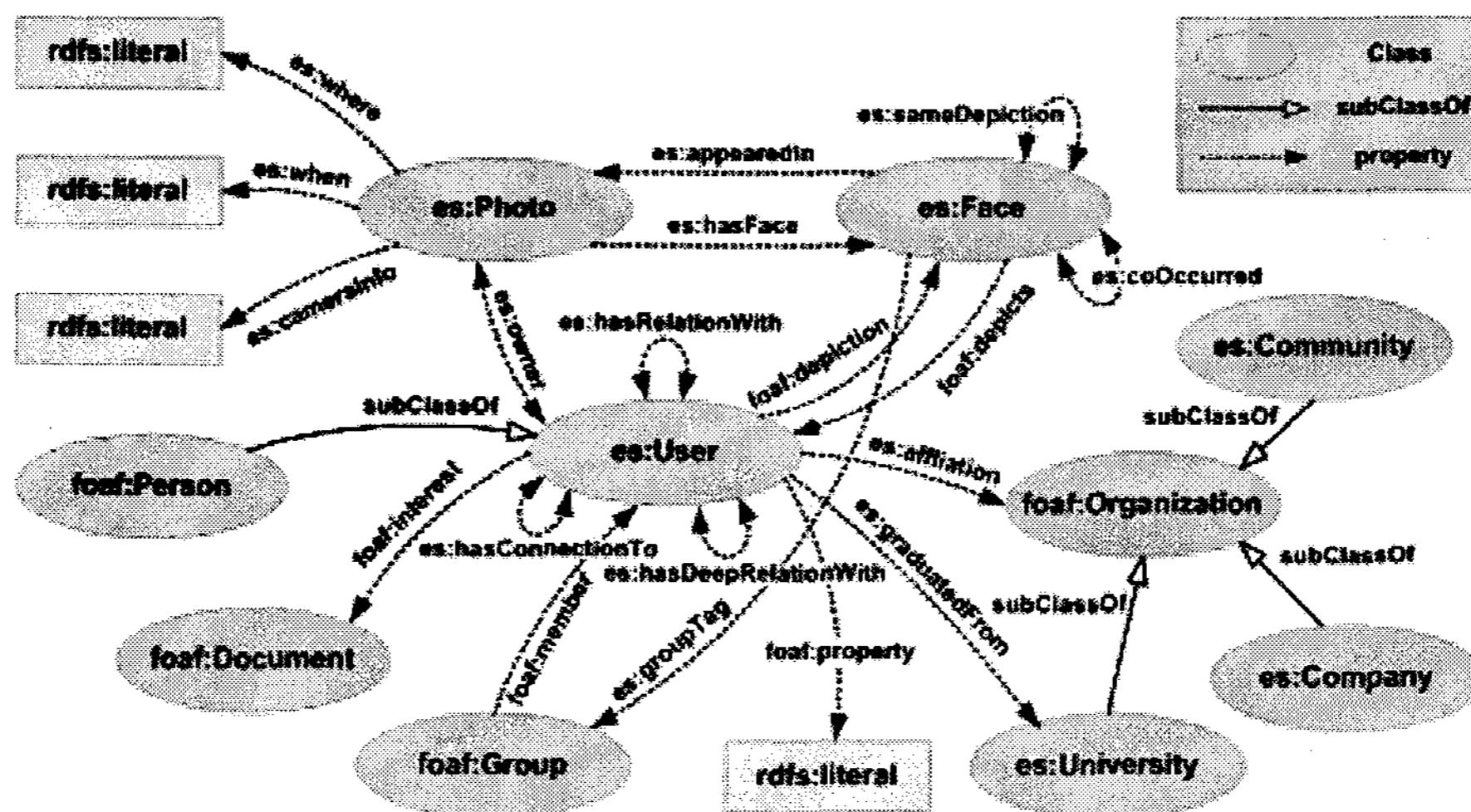


그림 2. 소셜 온톨로지

나타낸다.

본 논문에서 구축한 온톨로지는 특정 도메인의 지식 표현이 아닌 서로 다른 개체들을 어떻게 표현하여 의미적으로 연결할 수 있을 것인지를 초점을 두었다. 여기에서의 의미적 연결이란 개별 자원(resource)들에 명시적인 속성을 부여하여 연결하는 것을 말한다.

시맨틱 소셜 온톨로지의 구성은 FOAF에 정의된 클래스와 속성을 활용하였다. foaf:Person 클래스의 서브클래스로 es:User 클래스를 정의하여, 특정 사용자가 가지는 일련의 정보들을 FOAF에서 상속받은 속성과 새롭게 정의한 속성으로 표현하였다. 그리고 사람 간의 사회관계 및 사진에 관한 정보 그리고 얼굴 이미지에 관한 도메인을 따로 정의하여 온톨로지를 구축하였다.

그림 3은 사진(es:Photo), 얼굴 이미지(es:Face), 그리고 사용자(es:User)와의 관계를 나타낸다. 각각의 photo1, photo2에서 추출된 얼굴 이미지는 hasFace(역관계: appearedIn) 속성을 통해 Face 클래스의 얼굴 이미지들과 관계를 갖는다. photo1에서 추출된 얼굴 이미지 face1, face2, face3은 동시에 등장한 인물에 대한 정보를 정의한 coOccurred 속성에 의해 관계를 갖는다. Face 클래스의 인스턴스로 갖는다. 또한, owner 속성은 사진을 소유하고 있는 소유자가 누구인지에 관한 정보를 나타낸다. photo1, photo2에서 각각 추출된 face3과 face5는 sameDepiction 속성에 의해 동일한 인물을 묘사하고 있음을 알 수 있다. Face 클래스가 포함하고 있는 얼굴 이미지들에 관한 인스턴스들은 foaf:depicts(역관계: foaf:depiction) 속성에 의해 User 클래스의 인스턴스와 연결된다. 이와 같은 Photo, Face, User 클래스 간의 정의된 속성에 의해 사람 간의 소셜 관계를 발견한다.

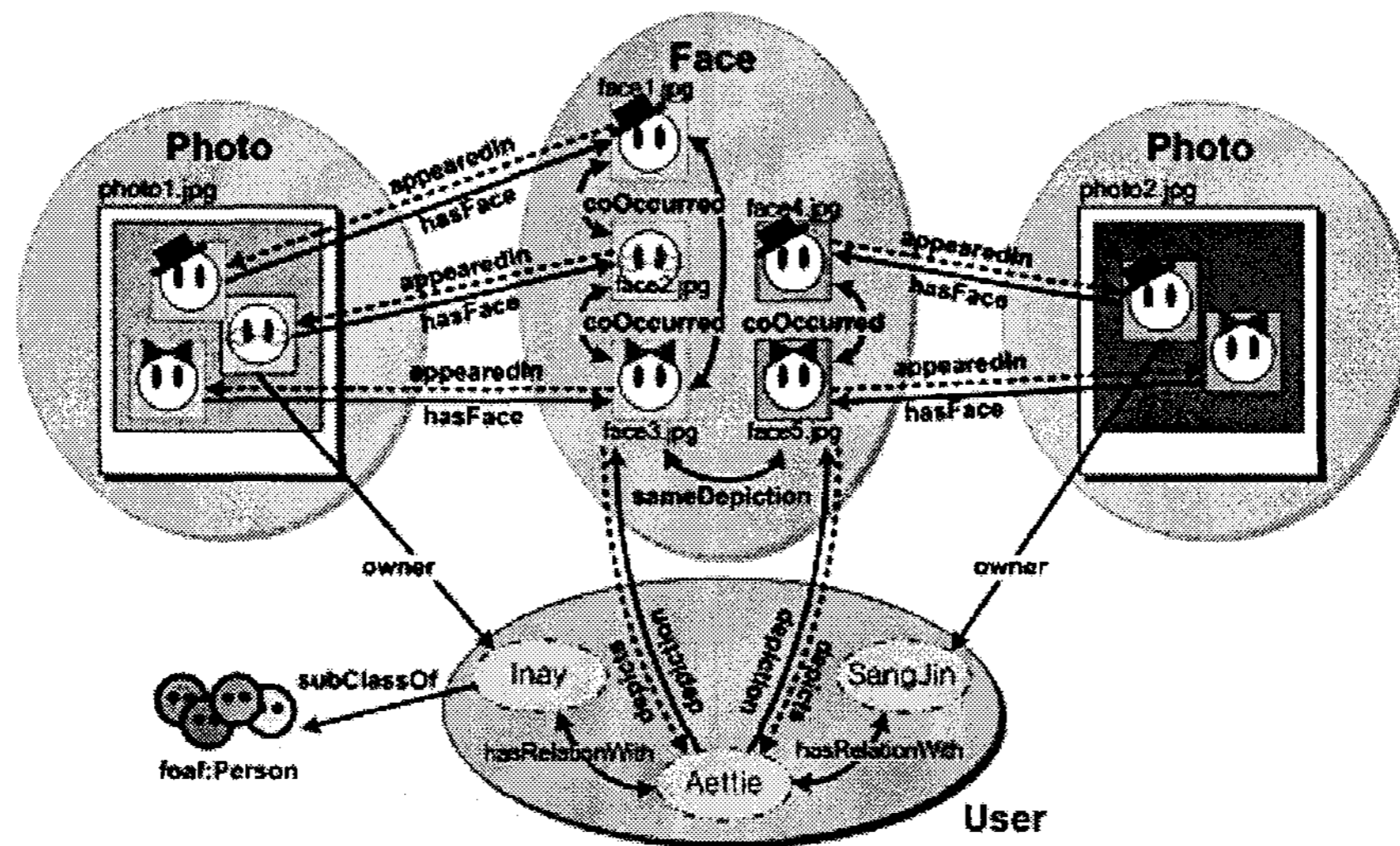


그림 3. Photo, Face, User 클래스 간의 관계

```
<es:User rdf:about="shlee">
  <foaf:name>Seunghoon Lee</foaf:name>
  <foaf:gender rdf:resource="male"/>
  <foaf:mbox
    rdf:resource="mailto:shlee@eslab.inha.ac.kr"/>
  <foaf:depicts
    rdf:resource="http://eslab.inha.ac.kr/seunghoon.jpg"/>
  <es:coWorker rdf:resource="Inay Ha"/>
  <es:hasFamily rdf:resource="Jonghee Lee"/>
  <es:hasConnectionTo rdf:resource="Sangjin Cha"/>
</es:User>
```

그림 4. RDF구문으로 표현한 개인 묘사

그림 4의 명세는 기계가 인식할 수 있는 RDF구문이다[13]. 아래의 내용은 다음과 같이 기술될 수 있다. es:User 클래스는 foaf:name, foaf:gender, foaf:mbox, foaf:depicts, es:coWorker, hasFamily, hasConnectionTo 속성으로 정의하여 개인의 이름, 성별, 이메일 주소, 대상의 이미지 묘사, 가족, 동료, 사회관계 등과 같은 정보를 표현한다.

3.3. 시맨틱 소셜 관계 분석

3.3.1. 소셜 관계 추론을 위한 도메인 규칙

사용자와의 연결 관계의 발견은 추론 엔진을 통해 이루어진다. 이러한 추론은 SWRL에 의해 정의된 도메인 규칙(domain rule)을 기반으로 수행하게 된다. SWRL(Semantic Web Rule Language)은 OWL의 하부 언어인 OWL DL 및 OWL Lite와 RuleML의 하부 언어인 Unary/Binary Datalog RuleML을 통합한 언어이다. SWRL은 유사 Horn 규칙(Horn-like rules)을 포함시켜 OWL의 공리(axiom)를 확장함으로써 유사 Horn 규칙을 OWL 지식 베이스와 통합시킬 수

있도록 하였다. SWRL의 규칙은 전제(antecedent (body))와 결과(consequent(head)) 간의 관계를 표시하는 형태를 갖는다. 이는 전제에 기술된 조건들이 만족될 때마다, 결과에 기술된 조건들도 만족된다는 의미를 갖는다. 다수의 원소들은 AND결합(conjunctive)으로 처리되며 $C(x)$, $P(x, y)$, $sameAs(x, y)$, $differentFrom(x, y)$ 등과 같은 형태로 표현된다. 표 1은 본 논문에서 정의한 몇 가지 규칙의 예이다.

표 1. 정의된 룰의 예

No	Domain rules
Rule 1	$Photo(?x) \wedge Owner(?x, ?y) \wedge hasFace(?x, ?z) \wedge depicts(?z, ?k) \wedge differentFrom(?y, ?k) \rightarrow hasRelationWith(?y, ?k)$ "Photo x 는 user y 가 소유하고 있고 facePhoto z 가 Photo x 에 등장하고 z 가 user k 를 묘사하고 y 와 k 가 동일인물이 아니면 hasRelation 의 개념을 갖는다."
Rule 2	$Face(?x) \wedge depicts(?x, ?y) \wedge depicts(?z, ?y) \wedge differentFrom(?x, ?z) \rightarrow sameDepiction(?x, ?z)$ "facePhoto x 가 y 를 묘사하고 facePhoto z 가 y 를 묘사하면 y 에 대한 sameDepiction 의 개념을 갖는다."
Rule 3	$User(?x) \wedge hasInterest(?x, ?y) \wedge hasInterest(?z, ?y) \wedge differentFrom(?x, ?z) \rightarrow SameInterest(?x, ?z)$ "facePhoto x 가 y 를 묘사하고 facePhoto z 가 y 를 묘사하면 y 에 대한 sameDepiction 의 개념을 갖는다."
Rule 4	$Face(?x) \wedge appearedIn(?x, ?y) \wedge foaf:depicts(?x, ?z) \wedge Owner(?y, ?k) \wedge FreqFace(?x, ?k) \rightarrow hasDeepRelationWith(?z, ?k)$ "facePhoto x 가 photo y 에 나타나고 x 가 z 를 묘사하고 k 가 y 를 소유하면 hasDeepRelationWith 의 개념을 갖는다."
Rule 5	$user(?x) \wedge hasRelationWith(?x, ?y) \wedge hasRelationWith(?y, ?x) \rightarrow connectivitionTo(?x, ?y)$ "User x 가 y 를 알고 있고 y 가 x 를 알고 있으면 서로 connectivity 의 개념을 갖는다."

3.3.2. 소셜 관계 추론

두 사용자의 의미적 관계 추론은 인스턴스의 사실(fact)과 정의된 도메인 규칙 기반으로 이루어진다. 제안된 시스템은 추론을 통하여 두 사용자 간의 친밀도가 높은 최상의 연결된 경로를 탐색하여 제공한다. 탐색 과정을 통해 발견해 낸 두 사용자 간의 최상의 연결 경로는 사용자와 밀접한 관계를 가지는 사람들을 통해 새로운 연결 관계를 추론해 냄으로써 사용자의 인맥 형성 네트워크를 증대시킬 수 있다.

추론을 기반으로 한 시맨틱 경로 탐색의 과정은 그림 5와 그림 6의 비교를 통하여 설명한다. 제안한 시스템의 최단 경로 탐색은 Floyd-Warshell[22] 알고리즘을 이용한다. 그림 5는 사진 소유자의 사진에 등장한 사람들(es:hasRelationWith) 간의 상호

관계를 방향성 있는 간선으로 연결하고 간선의 거리(distance)를 1로 설정하였다. 예를 들어, Floyd-Warshell 알고리즘으로 사용자 A에서 사용자 G까지의 최적의 경로를 탐색할 경우, 연결 노드의 수가 적은 사용자 C, 사용자 E를 거치는 경로가 선택된다. 하지만 노드 간의 연결 관계만을 고려하였기 때문에 사용자 간의 의미적 관계가 간과되었다.

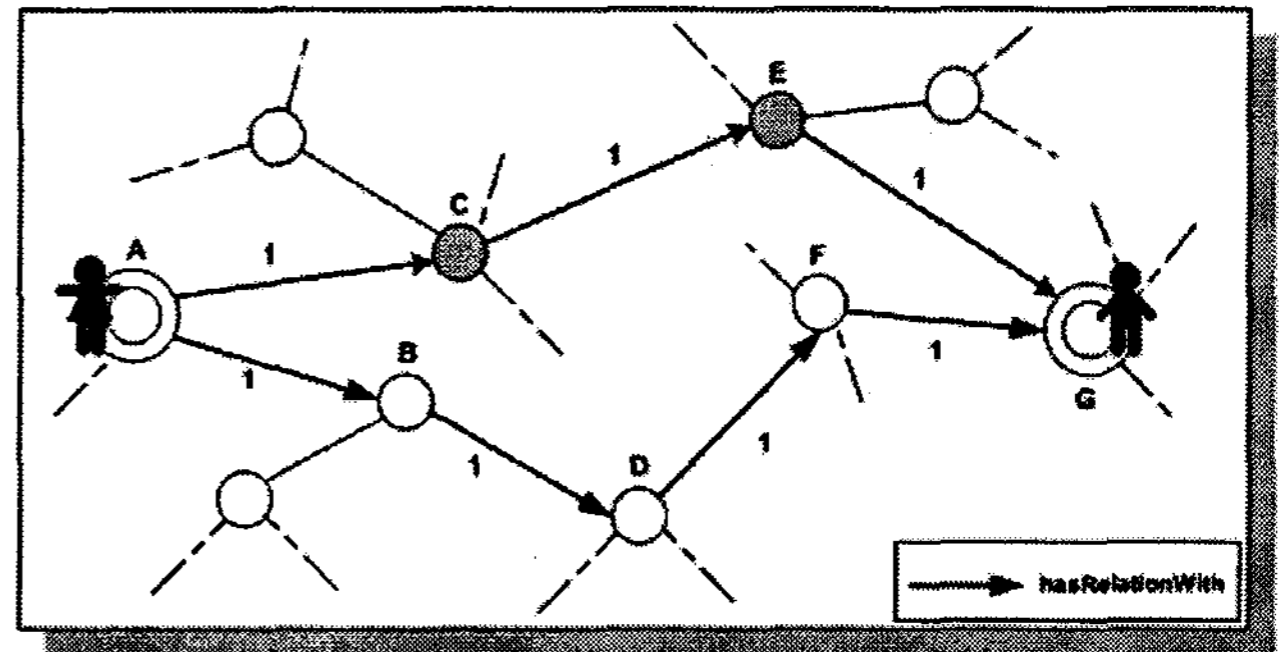


그림 5. 사진의 얼굴로부터 구성된 소셜 네트워크

이러한 문제를 보완하기 위해, 본 연구에서는 그림 6과 같이 최적의 경로를 탐색 과정에서 추론된 결과를 바탕으로 소셜 관계를 재구성 한다. 두 노드 사이의 거리는 의미적인 연결 관계를 고려하여 서로 다른 값을 부여한다. 그리고 새로 추가된 노드 간의 의미적 연결 관계는 가상의 연결(virtual link)을 갖게 되고, 이에 따라 새로운 가중치(거리)를 부여함으로써 소셜 그래프는 재구성된다.

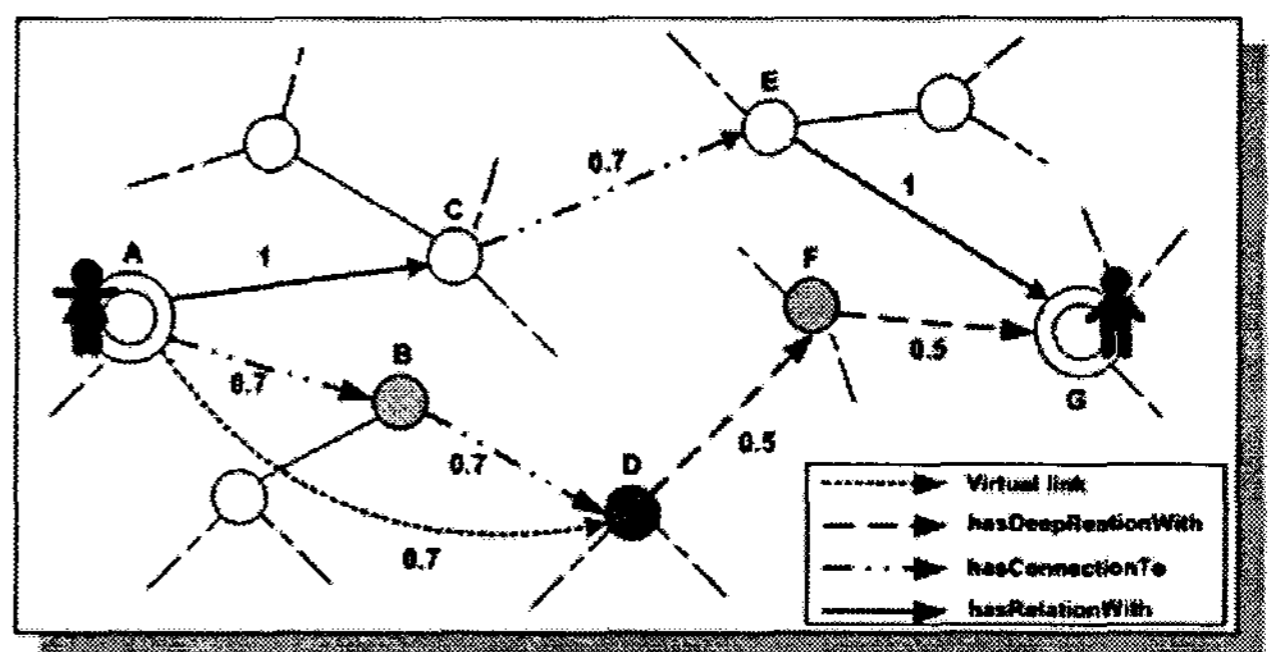


그림 6. 의미적 관계로 재구성된 소셜 네트워크

사용자 사이의 관계는 정의된 속성, es:hasRelationWith, es:hasConnectionTo, 그리고 es:hasDeepRelationWith에 의해 3단계로 나누어 부여한다. 예를 들어, 그림 6과 같이 사용자 A에서 사용자 G까지의 최적의 경로를 탐색할 경우, 사용자 A와 사용자 B가 각각 함께 찍은 사진을 서로 소유하고 있고 A가 같은 프로젝트에 참여하는 동료 관계(es:coWorker)일 경우, A와 B는 es:hasConnectionTo의 관계가 추론 된다. 마찬가지로 사용자 B와 사용자 D가 같은 프로젝트에 참여하는

동료 관계일 때, 사용자 A가 사용자 D의 직접적인 관계를 가지고 있지 않아도 추론에 의해 같은 프로젝트 동료임을 발견 할 수 있다. 또한 사용자 D가 소유한 사진에 사용자 F가 빈번하게 출현하였다면, 정의된 규칙에 의해 사용자 D와 사용자 F는 친밀한 관계(es::hasDeepRelationWith)임을 추론하게 된다. 이렇게 재구성된 소셜 그래프와 서로 다른 간선의 거리 값으로부터 Floyd-Warshall 알고리즘으로 최적의 경로를 탐색한다면, 그림 5의 결과와 다르게 의미적 관계가 부여된 A->B->D->F->G의 경로가 선택되어 진다.

4. 프로토타입

제안된 시스템의 프로토타입은 ATL Server 8.0 Framework 및 IIS6.0의 웹 서버 환경 하에, 얼굴 분석은 Intel의 컴퓨터 영상 프로그램 라이브러리인 OpenCV1.0¹²을 사용하였다. 추론 서버로 BEA Weblogic 9.2, 추론엔진은 Bossam 0.9b45¹³를 사용해서 이루어졌으며, Protégé3.3¹⁴ 플랫폼(platform)을 이용하여 온톨로지 및 도메인 규칙을 구축하였다.



그림 7. 웹 사용자 인터페이스

프로토타입 시스템의 사용자 인터페이스는 그림 7과 같은 4 가지 컴퍼넌트(component)로 구성된다. 우선 사용자가 사진을 업로드 할 수 있도록 하거나, 사용자가 소유하고 있는 사진으로부터 추출된 얼굴 이미지와 전체 사진을 보여주는 등 일반적인 기능의 컴퍼넌트가 제공된다. 사용자가 추출된 얼굴 이미지를 클릭하면 전체 사진 중 유사한 얼굴이 등장하는 사진을 검색하여 보여주고, 마지막으로 사용자가 네트워크상에서 소셜 관계를 가지는 사람들을 그래프로 표현하고, 사람 간의 최적 경로를 탐색하는 기능을 제공한다.

그림 8은 얼굴 이미지를 클릭하였을 때 Face 클래스에 정의된 속성을 기반으로 선택된 얼굴

이미지와 관련이 있는 전체 사진을 검색하여 사용자에게 보여주는 예이다.



그림 8. 얼굴 개체를 포함한 전체 이미지 목록 나열

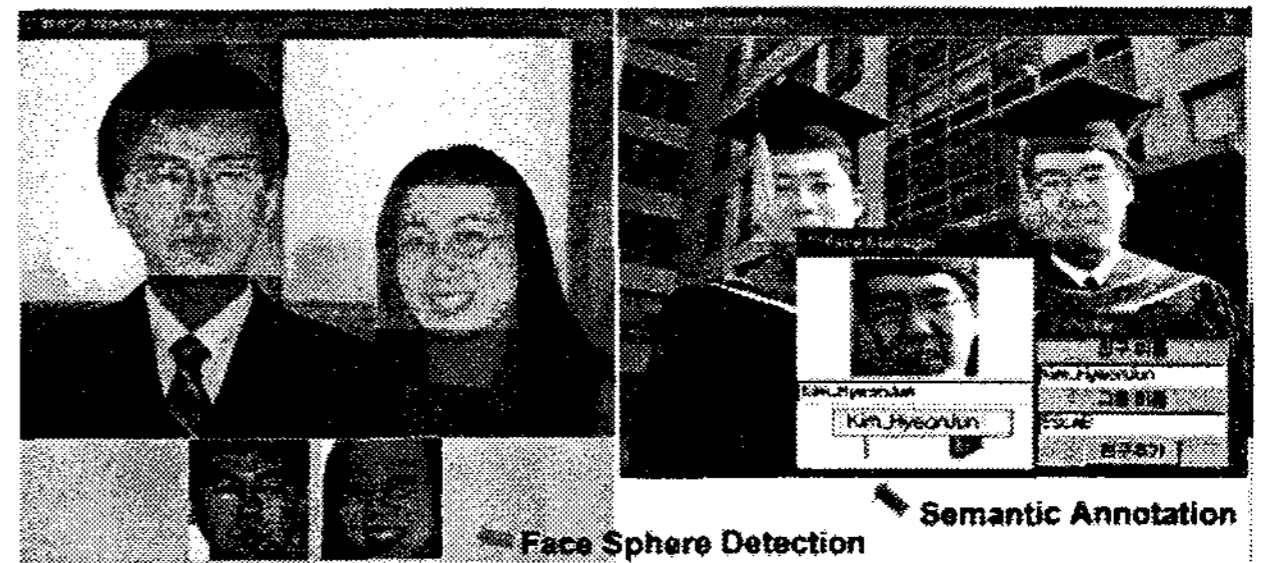


그림 9. 얼굴 영역 인식 및 의미적 주석화

그림 9는 사용자가 사진을 업로드 하였을 때 얼굴 영역을 인식하고 추출된 얼굴 개체를 정의된 User 클래스의 사용자와 연결하는 부분을 나타낸다. 즉, 얼굴 사진은 온톨로지 내의 사용자 프로파일과 매칭되어 식별이 가능한 의미적 정보로 표현된다.

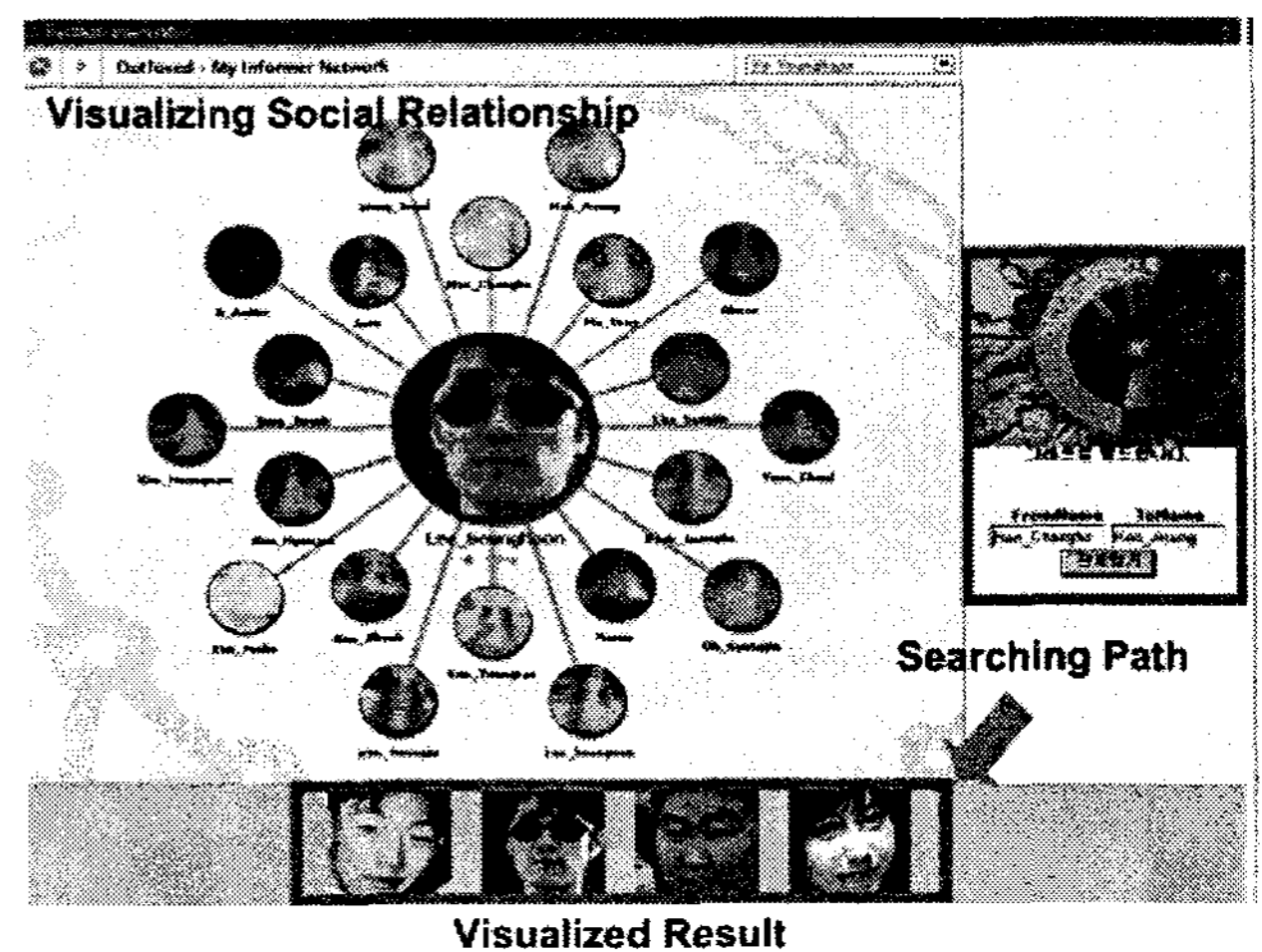


그림 10. 소셜 관계 그래프 및 검색

마지막으로 시스템은 그림 10과 같이 사용자의 사진에 등장하는 사람들을 기반으로 그 관계를 소셜 네트워크 그래프로 시각화 한다. 특정 개인과의

¹² <http://www.intel.com/technology/computing/opencv/>

¹³ <http://bossam.wordpress.com/>

¹⁴ <http://protege.stanford.edu/>

소셜 관계를 찾고자 질의를 하면 시스템은 온톨로지에 정의된 클래스와 속성, 그리고 도메인 규칙을 적용하여 사용자 간의 정보적 연결성을 추론하여 제공한다.

5. 실험 및 결과

5.1. 실험 평가 및 데이터 분석

제안하는 시스템을 이용하여 발견된 두 사용자간의 연결 경로가 실제로 높은 친밀도를 가지는 인맥들을 바탕으로 탐색되어졌는가를 실험을 통하여 평가하였다.

5.1.1. 데이터 집합

성능 평가를 위해 30명의 사용자(사진 소유자)가 직접 사진들을 업로드 하여 총 307장의 사진 데이터를 구축하였다. 그리고 업로드 된 사진으로부터 추출된 378개의 얼굴 이미지를 이용하였다. 여기에서 추출된 얼굴 이미지의 개수는 온톨로지 내에 클래스와 속성으로 정의된 사용자 프로파일의 개수와 동일하다고 가정하였다. 실험을 위한 데이터 집합은 표 2와 같다.

표 2. 데이터 집합

# of user profile	# of photo owner	# of photos
378	30	307

cf. number of face photos = number of user profiles

5.1.2. 실험 평가 방법

사진을 실제 소유하고 있는 사용자 30명 중 한 명과 시스템에서 임의의 사용자 한 명을 선택하여 이 두 사람 간의 소셜 연결 경로를 추론하고 그 결과가 얼마나 정확하게 예측 되었는가를 평가한다. 시스템이 표 3과 같은 소셜 연결 경로 질의 50개를 수행하도록 하였고, 질의된 두 사람의 연결 관계는 실제 두 단계 이상, 다섯 단계를 넘지 않는 대상으로 선택하였다.

표 3. 소셜 관계 분석을 위한 질의의 예

Query examples
Searching the best linked-path between "Seung-Hoon Lee and "Young-sook Kim"

추론 결과를 바탕으로 탐색된 연결 경로(SR)의 정확성은 사용자 간의 Binary Relationship을 바탕으로 탐색된 경로(BR)와의 비교, 분석을 통하여 평가된다.

3.3절에서 언급하였듯이 소셜 그래프에서 각각의

BR과 SR은 Floyed-Warshell[22] 알고리즘을 이용하여 최단 경로를 탐색한다. BR은 foaf:knows 속성과 같은 개체들 간의 연결 관계를 방향성 간선으로 표현하고 노드 간의 연결 관계가 있으면 그 거리를 1로 설정하였다. 여기서 사진 소유자의 사진에 등장한 사람은 foaf:knows 속성을 갖는다고 가정하였다. 제안된 방법(SR)은 온톨로지에 정의된 클래스, 속성 및 규칙에 의해 추론된 결과를 반영하였다. 새로 추론된 노드 간의 의미적인 연결 관계는 소셜 그래프로부터 가상의 연결(virtual link)을 갖게 된다. 사용자 간의 추론된 관계 및 가상의 연결 관계에 대하여 연결 간선의 거리를 3단계로 나누었다. 예를 들어, 두 사용자의 관계가 es:hasRelationship일 경우의 간선 거리를 1, es:hasConnectionTo일 경우는 0.7, 그리고 es:hasDeepRelationship일 경우는 0.5로 각각 추론된 결과를 기반으로 부여하였다.

제안된 방법의 평가는 질의에 대한 결과로 탐색된 두 사람 간의 최적 경로가 실제 사용자와의 소셜 관계와 얼마나 유사한가를 기준으로 한다. 이는 테스트 데이터 집합에 대한 매칭 데이터의 비율로 판단하며, 계산식은 다음과 같다.

$$hit - rate = \frac{|Test \cap Result|}{|Test|}$$

Test: 테스트 데이터 집합, Result: 탐색된 경로 결과 집합

5.2. 실험 및 결과

그림 11은 BR과 SR에 대한 적중률을 비교한 것이다. 그래프에서 보는 바와 같이 제안된 시스템에서 탐색된 소셜 연결 경로 SR의 적중률은 72%로 BR의 적중률 56%보다 16% 향상된 성능을 보였다. 즉, 제안한 시스템이 친밀도가 높은 사람 간의 최적 연결 경로를 탐색하는 데 효율적이라는 것을 알 수 있다.

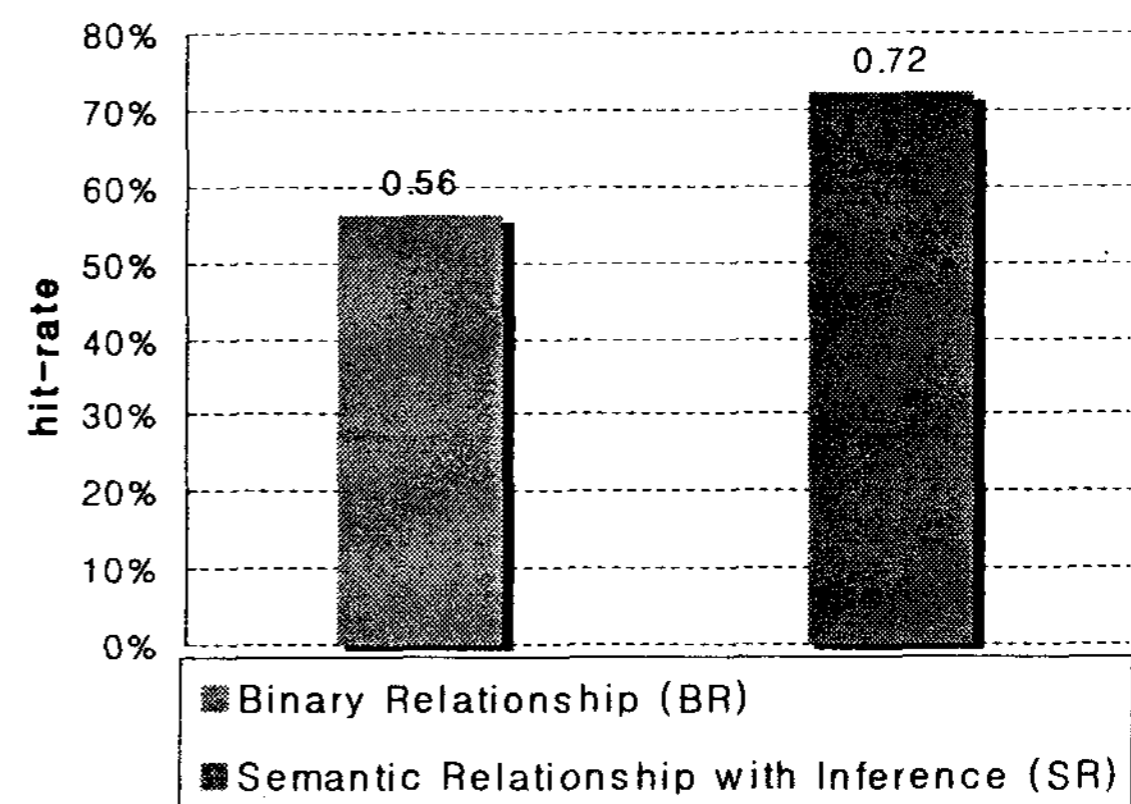


그림 11. 적중률 비교 결과

BR에서는 사진에 나타난 사용자의 연결 관계(es:hasRelationship) 만을 경로 탐색에 사용하여, 각 노드(사용자)가 어떤 의미로 연결되어졌는지

파악할 수 없었다. 또한 모두 같은 연결 관계의 간선 거리로 최단 경로를 탐색하였기 때문에, 단순히 연결된 노드의 개수가 적어 그 거리가 짧은 경로를 결과로 제공하였다. 이에 비해 제안된 SR 기반에서는 온톨로지를 바탕으로 사진과 사용자, 사용자와 사용자 간의 관계를 명확히 하고 정의된 규칙에 의해 새로운 관계 및 지식을 추론함으로써, 새로운 가상의 연결 관계를 생성할 수 있었다. 또한 두 사용자 간의 관계를 의미적으로 추론하여 그 간선 거리를 상이하게 설정할 수 있었다. 추론된 새로운 연결 관계와 의미에 따라 다르게 부여된 간선 거리를 바탕으로 한 최단 경로 탐색은 보다 정확한 두 사용자의 간의 의미적인 연결 경로를 제공할 수 있었다.

하지만 실험에 참가한 커뮤니티 사용자(사진 소유자)가 협소하게 구성이 되어있어, 서로 복잡하게 연결되어 있는 집단(대체로 사진 소유자들)과 느슨하게 연결되어 있는 집단(얼굴만 추출된 사용자들)이 확연하게 구분되어 있었다. 그로인해 연결 관계를 많이 가지고 있지 않은 일반 사용자까지 다양한 연결 경로를 가지고 있지 못하였다. 더불어 소셜 관계에 대한 도메인 규칙이 풍부하고 세부적으로 정의되어 있지 않으면 새로운 관계 추론의 한계가 있었다. 따라서 복잡한 네트워크 내에 소셜 관계를 추론 가능하도록 명확하고 새로운 규칙들을 많이 정의해야만 향상된 성능을 갖게 될 것이다. 표 4는 일부 사용자 간의 관계 추론의 결과이다.

표 4 사용자 간의 관계 추론의 결과

User	Lee_Seunghoon	Han_Changho	Kim_Youngju	Cha_Sangjin	Ha_inay	Kim_Youngsuk	Kim_Youngtae	Lee_Hokyung	Kim_Hyunjun	Jung_Jina	Choi_Hyunhee
Lee_Seunghoon		C	R		D		C		D		R
Han_Changho			D	C		R					
Kim_Youngju		D		C	R		C		R		
Cha_Sangjin					D		C	D			
Ha_inay			R	D		C		D	D		R
Kim_Youngsuk	R				C						C
Kim_Youngtae	C		C	R					R		
Lee_Hokyung				D		R			C		R
Kim_Hyunjun	D				D			C			R

6. 결론 및 향후 연구

기존의 소셜 네트워크 서비스가 사회적 개체들 간의 관계에 대하여 명확하게 정의되어 있지 않아 의미적 관계를 발견 해 내는데 한계가 있다. 이를 해결하기 위해 본 연구에서는 시맨틱 웹 기반의 소셜

네트워크 분석 시스템을 제안하여, 사용자 개인의 사진으로부터 소셜 네트워크를 형성하고 시각화하는 프로토타입 시스템을 구현하였다. 즉, 개인 사용자들이 업로드한 사진들을 분석하여 의미적인 가상 커뮤니티를 반자동적으로 구성하고 새로운 관계 추론에 활용하였다.

제안된 시스템은 기존의 소셜 네트워크 서비스와는 다르게 의미적인 관계의 추론을 위한 온톨로지 및 도메인 규칙을 정의함으로써 보다 향상된 소셜 관계를 제공해 준다. 또한 정보자원을 컴퓨터가 처리할 수 있는 OWL, RDF/RDFS, 그리고 SWRL을 사용하여 의미적 정보로 표현하고 연결함으로써 상이한 시스템 간의 재사용 및 공유의 기반을 마련하였다. 하지만, 복잡하게 표현될 수밖에 없는 개인이나 사회적 관계를 상세하게 정의하는 데에 한계가 있었다. 유연한 사회적 관계를 상세화하고 보다 다양한 의미적 관계를 추론하기 위해 체계적인 도메인 규칙의 확장이 필요하다. 그리고 사용자들의 개인 프라이버시 문제도 고려해야 한다.

지금까지 소셜 네트워크 서비스는 대부분 젊은 층과 디지털 사회에서 활발히 활동하고 있는 사람들에게 국한되어 있었다. 향후 학교, 기업 관련 소셜 네트워크 서비스나 교육 관련 서비스, 그리고 개인 사진 추천 시스템 등과 같은 다양한 응용 서비스에 적용하여 제안된 시맨틱 소셜 네트워크 시스템을 활용하는 방안을 모색해야겠다.

참고문헌

- [1] Garton, L., Haythornthwaite, C. and Wellman, B.: Studying Online Social Networks. Journal of Computer-Mediated Communication, Vol. 3, No. 1 (2000)
- [2] Wennerberg, P. O.: Ontology Based Knowledge Discovery in Social Networks, Final Report, JRC Joint Research Center (2005)
- [3] Aleman-Meza, B., Nagarajan, M., Ramakrishnan, C., Ding, L., Kolari, P., Sheth, A. P., Arpinar, I. B., Joshi, A. and Finin, T.: Semantic Analytics on Social Networks: Experiences in Addressing the Problem of Conflict of Interest Detection, In Proceedings of the 15th International Conference on WWW, pp. 407-416 (2006)
- [4] Berkowitz, S. D.: An Introduction to Structural Analysis: The Network Approach to Social Research. Butterworth, Toronto (1982)
- [5] Wasserman, S. and Faust, K.: Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge (1994)
- [6] Chen, C.: Visualizing Semantic Spaces and Author Co-Citation Networks in Digital Libraries, In Information Processing Management, Vol. 35, No. 3, pp.

401-420 (1999)

- [7] Barabási, A.-L.: *Linked: The New Science of Networks*. Perseus Publishing, Cambridge, MA (2002)
- [8] Nascimento, M. A., Sander, J. and Pound, J.: Analysis of SIGMOD's Co-Authorship Graph. *SIGMOD Record*, Vol. 32, No. 3 (2003)
- [9] Mika, P.: *Social Networks and the Semantic Web: Semantic Web and Beyond*, Springer (2007)
- [10] Finin, T., Ding, L., Zhou, L. and Joshi, A.: Social Networking on the Semantic Web, In *The Learning Organization*, Vol. 12, No. 5, pp. 418-435 (2005)
- [11] Ghita, S., Nejdl, W. and R. Paiu.: Semantically Rich Recommendations in Social Networks for Sharing and Exchanging Semantic Context, In *Proceedings of the Ontologies in P2P Communities Workshop at ESWC (2005)*
- [12] <http://www.foaf-project.org/2004/us/about.html>
- [13] <http://xmlns.com/foaf/0.1/>
- [14] Ding, L., Zhou, L., Finin, T. and Joshi, A.: How the Semantic Web is Being Used: An Analysis of FOAF, In *Proceedings of the 38th International Conference on System Sciences (2005)*
- [15] Golbeck, J., Parsia, B. and Hendler, J.: Trust Networks on the Semantic Web, In *Proceedings of the 12th International Conference on WWW (2003)*
- [16] Grimnes, G. A., Edwards, P. and Preece, A.: Learning Meta-descriptions of the FOAF network, In *Proceedings of the 3rd International Semantic Web Conference*, pp. 152-165 (2004)
- [17] Ounnas, A., Davis, H. C and Millard, D. E.: Semantic Modeling for Group Formation, In *Proceedings of Workshop on Personalization in E-Learning Environments at Individual and Group Level at UM2007 (2007)*
- [18] Kustanowitz, J. and Shneiderman, B.: Motivating Annotation for Personal Digital Photo Libraries: Lowering Barriers while Raising Incentives. Technical Report, University of Maryland, USA (2005)
- [19] Biddulph, M.: A Semantic Web Shoebox - Annotating Photos with RSS and RDF, In *Proceedings of XML Europe 2003 Conference (2003)*
- [20] Axelrod, A., Golbeck, J. and Shneiderman, B.: Generating and Querying Semantic Web Environments for Photo Libraries, Technical Report, University of Maryland, USA (2005)
- [21] Shevade, B., Sundaram, H. and Xie, L.: Modeling Personal and Social Network Context for Event Annotation in Images, In *Proceedings of the Joint Conference on Digital Libraries (2007)*
- [22] Cormen, T. H., Leiserson, C. E., Rivest, R. L. and Stein, C.: *Introduction to Algorithms*, 2nd Ed., The MIT Press (2001)
- [23] SWRL <http://www.w3.org/Submission/SWRL/>