

개인화서비스에서 시맨틱웹 기반의 사용자 선호정보 공유에 관한 연구

김주연[†], 김종우^{**}, 김창수^{***}

요 약

사용자의 요구와 선호도에 따라 적합한 정보를 제공해주는 개인화서비스에 대한 많은 연구와 개발이 진행되어왔다. 하지만 기존의 연구들은 단일시스템 내에서 사용자의 선호정보를 관리하기 때문에, 다양한 서비스들 사이에서 이러한 정보를 공유하기 어렵다는 문제가 있다. 본 논문에서는 이러한 문제를 해결하기 위한 하나의 방법으로 시맨틱웹 기반의 사용자 선호정보 공유 모델을 제안한다. 이 모델은 사용자의 선호정보가 각 서비스의 특성을 반영하는 서비스 온톨로지 상에서 의미를 기반으로 기술되어지고 공유되어지도록 한다. 또한, 본 논문에서는 제안모델을 지원하는 미들웨어 구현을 통해 제안모델을 분석하고 평가한다. 제안하는 접근방법은 각 서비스를 중심으로 사용자 선호정보를 기술할 수 있으면서도 다양한 서비스들 간에 그 정보를 공유할 수 있기 때문에, 기존의 연구보다 효율적인 개인화서비스를 제공할 수 있다는 장점이 있다.

A Study on User Preference Sharing based on Semantic Web in Personalized Services

Ju-Yeon Kim[†], Jong-Woo Kim^{**}, Chang-soo Kim^{***}

ABSTRACT

Many personalized Services that provide users with adaptive information according to users' requirements and preferences have been researched and developed. However, existing approaches are difficult to share a user's information among heterogeneous services because these approaches manage users' preferences in a single system. In this paper, we propose a user preference sharing model based on the Semantic Web as a solution to resolve the problem. Our model enables user preferences to be described and shared over service-specific ontologies which are affected by the feature of each service. Our model is analyzed and evaluated with an implementation of the middleware that supports our model. Our approach has the advantage of providing more efficient personalized services than existing approaches because it can describe users' preferences centering around each service and share these information among heterogeneous personalized services.

Key words: Personalized Service(개인화 서비스), User Preference Sharing(사용자 선호정보 공유), Semantic Web(시맨틱웹)

※ 교신저자(Corresponding Author): 김창수, 주소: 부산광역시 남구 대연 3동 599-1(608-737), 전화: 051)620-6394, FAX: 051)620-6394, E-mail: cskim@pknu.ac.kr
접수일: 2007년 7월 19일, 완료일: 2007년 9월 5일
[†] 준회원, 부경대학교 정보보호학과

(E-mail: jyim@pknu.ac.kr)
^{**} 준회원, 대구대학교 정보통신공학부
(E-mail: jkim@daegu.ac.kr)
^{***} 종신회원, 부경대학교 정보보호학과

1. 서 론

인터넷의 발전과 함께 급증하고 있는 정보와 다양한 서비스들 사이에서 사용자가 자신이 원하는 정보를 검색하는 것이 점점 더 어려워지고 있다. 이에 사용자의 요구와 선호도(preferences)에 따라 적합한 정보를 제공해주는 개인화서비스(Personalized Services)에 대한 많은 연구와 개발이 진행되어왔다 [1,2]. 하지만, 기존의 연구들은 주로 단일 시스템 내에서 해당 서비스의 특성에 따라 사용자의 선호정보를 관리하기 때문에, 다양한 서비스들 사이에서 이러한 정보를 공유하기 어렵다. 즉, 하나의 서비스에서 사용자가 등록한 선호정보는 다른 서비스에서는 일반적으로 반영되지 못한다. 따라서 이러한 사용자의 선호정보를 서로 다른 서비스 사이에서 교환할 수 있는 방법이 요구된다. 특히, 최근 관심이 고조되고 있는 유비쿼터스 환경에서는 주로 제한된 성능, 제한된 사용자 인터페이스 및 낮은 대역폭의 통신환경을 제공하는 단말기에서 분산된 다양한 서비스에 대한 접근이 많이 이루어지기 때문에, 개인화서비스 뿐 아니라 다양한 서비스들 사이에 사용자 선호정보의 공유를 통한 효율적인 개인화서비스 지원이 요구될 것이다.

이러한 사용자 정보 공유 문제를 해결하기 위해 사용자 정보를 교환하기 위한 표준들이 다양한 분야에서 연구되고 있다. 하지만 이러한 접근방법들 역시 사용자 선호정보에 대한 의미적인 공유가 이루어지지 않고 표준에서 규정하고 있는 제약사항이 존재하기 때문에, 다양한 서비스들에 존재하는 용어와 의미를 완전하게 반영하지 못하며 다른 시스템에 존재하는 이음동의어나 유사한 의미에 대해서는 서로 공유하지 못한다.

이에 본 논문에서는 이러한 문제를 해결하기 위한 하나의 방법으로 시맨틱웹 기반의 사용자 선호정보 공유 모델을 제안한다. 시맨틱웹은 컴퓨터가 웹상의 자원(resource)을 의미기반으로 기술(describe)하고 처리할 수 있도록 하는 기술로서, 기하급수적으로 증가하고 있는 웹상의 자원을 능동적으로 검색할 수 있도록 기반구조를 제공하는 차세대 웹 기술이다 [3,4]. 이러한 시맨틱웹 기술은 서로 다른 정보와 형식의 콘텐츠를 제공하는 서비스 제공자간에 사용자의 요구사항이나 선호정보를 의미적으로 공유할 수

있는 기반을 제공할 수 있다. 따라서 본 연구는 서비스 온톨로지(Service-specific Ontology)들 상에서 사용자의 선호정보를 기술함으로써 사용자의 선호정보를 그 의미를 기반으로 정확하게 표현(representation)할 수 있는 방법을 제공하고, 다양한 개인화 서비스들 사이에서 사용자 선호정보의 의미적 분석을 통해 강력한 상호운용성(interoperability)을 제공하는 것을 목표로 한다.

본 논문의 구성은 다음과 같다. 2장에서는 개인화 서비스 연구동향과 시맨틱웹 기술에 대해 분석하고, 3장에서는 제안하는 사용자 선호정보 공유 모델의 접근방법과 요소기술들에 대해 기술하며, 4장에서는 제안모델을 분석 및 평가한다. 5장에서는 본 논문을 결론짓고, 향후연구를 제시한다.

2. 관련연구

2.1 개인화 서비스 연구동향

인터넷의 급속한 발전에 따라 사용자에게 제공되는 정보들은 기하급수적으로 증가하고 있으며, 이 수많은 정보들 중에서 사용자가 자신에게 필요한 정보를 찾는다는 것은 사용자에게 큰 부담을 부가하게 되었다. 이에 최근에 사용자들의 요구와 선호도에 따라 정보를 분석하여 적합한 정보를 제공해주는 개인화서비스에 대한 많은 연구와 개발이 진행되고 있다 [5-8]. 사용자의 선호도를 고려하여 비슷한 사용자의 성향에 맞는 정보를 제공해주는 서비스는 현재 아마존(Amazon.com), 옥션(auction.co.kr)등 전자상거래 분야에서의 성공적인 개인화서비스 사례로 제시되고 있으며[2], 국내에서도 모바일상에서 개인화된 서비스로 SKT의 Imm Expert Agent 서비스가 사용자의 과거 이용내역을 토대로 뉴스, 날씨, TV, 영화, 맛집 등의 8가지 아이টে에 대해 인기 있는 정보를 제공하고 있다.

이러한 개인화서비스를 지원하기 위해서는 서비스를 제공하는 시스템들이 개인의 요구하는 정보, 사용자의 행동, 선호도등을 분석하고 그에 적합한 콘텐츠나 서비스를 제공해야한다[1,2]. 이러한 개인화서비스를 효율적으로 제공하기 위해서는 우선적으로 사용자의 관심 정보가 효율적으로 수집되고 제공되어야 하기 때문에, 사용자 프로파일링 기술에 대한 많은 연구들이 다양한 활용분야에서 진행되고

있다.

Widyantoro[5]는 기계학습방법(mechine learning)을 통해 사용자의 행동을 관찰하고 교육(tranning)을 통해서 사용자의 관심있는 정보를 모델링하고 관심있는 항목(supervised learning)과 관심없는 항목(unsupervised learning)을 사용자의 프로파일링에 사용하였고, Wong[6]은 바이시안 네트워크(Bayesian network)를 이용하여 각 영역에 대한 확률값을 이용하는 방법을 제안하였다. 개인화서비스가 응용분야에 활용된 예로서, MUSIC[7] 시스템은 사용자의 현재 감정이나 선호하는 음악장르 등을 수집하여 데이터베이스에 저장하여 음악분야에 관한 개인에게 적합한 서비스를 제공하고 있으며 LBS(Location-Based Services)에서의 개인화서비스[8,9] 연구에서는 개인의 선호정보를 데이터베이스 기반이나 XML 기반으로 프로파일링하여 웹상에서 개인화 서비스로 활용 가능하도록 하였다.

이러한 연구들은 개인화된 서비스를 제공하기 위해서 단일시스템 내에서 사용자 프로파일을 통해 사용자의 요구나 선호정보 등을 생성하고 사용한다. 하지만 최근 개인화서비스를 제공하는 서비스들이 증가하면서, 다양한 서비스들 사이에 사용자 프로파일 공유에 대한 문제가 대두되고 있다. 따라서 이러한 문제들의 해결을 위해 서로 다른 시스템들 간에 정보들을 교환하기 위한 일반화된 기준이 요구되었으며, 서비스 도메인별로 표준화가 진행되고 있다. 예를 들어, 교육분야에서는 IEEE의 PAPI[10]와 IMS의 LIP[11]와 같은 표준들을 제시하고 이를 활용하기 위한 노력들이 이루어지고 있다. 하지만 이러한 노력에도 불구하고, 다양한 서비스들 사이에 개인화서비스 지원을 위한 사용자 정보의 공유에 있어서 다음과 같은 문제점들이 여전히 존재한다. 먼저, 표준에서 규정하고 있는 제약사항에 따라 사용자의 선호정보를 표현하거나 교환하는 것은 다양한 서비스들에 존재하는 용어와 의미를 완전하게 반영하지 못한다. 또한, 단일시스템에 존재하는 사용자의 정보는 그 서비스를 위해 설계된 스키마나 용어에 의해 저장되고 처리되기 때문에 다른 시스템에 존재하는 동일한 사용자의 유사한 의미나 다른 용어로 정의된 정보(이음동어) 및 동음이의어에 대해 서로 공유하지 못한다는 문제점이 있다.

2.2 시맨틱웹

시맨틱웹은 컴퓨터가 웹상의 자원(resource)을 의미 기반으로 표현하고 처리할 수 있도록 하는 기술로서, 기하급수적으로 증가하고 있는 웹상의 자원을 지능적으로 검색할 수 있도록 기반구조를 제공하는 차세대 웹 기술이다[3,4]. 시맨틱웹 기술들은 1990년대 말에 Tim Berners Lee를 중심으로 시맨틱웹 비전을 발표하면서 관심이 고조되어 왔으며, 최근에는 교육분야[12], 의학분야[13], 관광분야[14] 등 다양한 분야에서 시맨틱웹을 활용하고자 하는 연구들이 활발히 진행되고 있다.

시맨틱웹은 시맨틱웹 기술 계층(그림 1)의 각 요소기술들에 의해 웹 상의 리소스에 대한 의미적인 표현과 처리를 제공한다. 시맨틱 웹의 가장 기본이 되는 URI(Uniform Resource Identifier)는 웹 자원의 고유 식별자로 사용되며, XML + Namespace + XMLSchema는 리소스의 표현을 구조화할 수 있도록 지원함으로써 문서의 구조적인 측면에서 상호용성을 제공한다[14]. RDF + RDFSchema는 웹상의 자원(Resource)들을 정의하고 기술하기 위한 XML 기반의 데이터 모델로서, 컴퓨터가 리소스를 의미적으로 처리할 수 있는 기반을 제공한다. Ontology는 웹 환경에서 자원에 대한 개념과 관계를 정의함으로써 웹상의 자원에 대한 의미를 정의한다. 이렇게 기술된 웹상의 자원들은 Logic, Proof, Trust 계층의 요소기술들을 통해 자원에 대한 지능화된 접근과 신뢰성을 제공받게 된다[10].

특히, RDF와 Ontology는 시맨틱웹의 핵심적인 기반구조로서의 역할을 담당한다. RDF(Resource Description Framework)는 웹상의 자원들을 정의하

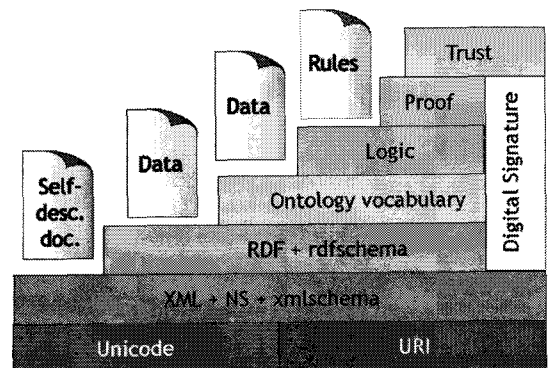


그림 1. 시맨틱웹 기술 계층

기 위한 XML기반의 데이터 모델로서, RDF에서는 데이터를 정의하기 위해 자원간의 관계를 Triple 구조(subject-predicate-object)로써 표현함으로써 컴퓨터가 의미를 처리할 수 있는 기반을 제공한다. RDF는 데이터모델을 컴퓨터가 처리하기 위해서 RDFS(RDF Schema)에 따라 인코딩하게 된다 [15,16].

온톨로지(Ontology)[3]는 분산된 웹 환경의 자원 간에 의미적인 연관성과 관계를 정의한다. 시맨틱 웹에서는 웹 환경에 존재하는 모든 자원을 정의하기 위해서 개념(concept 또는 class)들을 사용하고 있다. 이렇게 개념으로 정의된 클래스들은 그 의미적인 처리를 위해서 그 개념간의 관계를 속성(property)으로 정의하고 있다. 이 속성으로 정의된 관계들은 컴퓨터가 웹 자원들을 의미적으로 이해하고 처리할 수 있는 기반이 된다. 즉, 프로퍼티에서는 개념간에 무엇을 정의하고 있는지, 어떤 범위의 개념들을 포함하는지 등의 다양한 관계 등을 컴퓨터가 처리 가능하도록 규정하는데 사용할 수 있다. 온톨로지를 정의하기 위한 여러 가지 방법이 있지만, 특히 OWL(Web Ontology Language)[17]은 W3C에서 웹 자원의 관계를 풍부하게 정의할 수 있도록 제공하는 온톨로지를 기술하기 위한 언어이다. OWL은 해당 도메인에 존재하는 개념간의 관계, 제약사항등을 다양하게 정의할 수 있기 때문에 의미적인 처리 및 도메인의 지식에 대한 풍부한 표현이 가능하게 한다.

3. 사용자 선호정보 공유시스템

3.1 사용자 선호정보 공유 모델

분산 환경에서 다른 시스템의 자원을 활용하기 위해서는 시스템 또는 콘텐츠간의 상호운용성(interoperability)이 필요하게 된다. 상호운용성이란, 시스템 또는 데이터가 사람의 특별한 노력 없이도 다른 시스템이나 데이터와 함께 잘 동작될 수 있는 능력으로 이기종 시스템들 간의 정보의 통합이나 교환을 말한다. 본 연구의 목적은 다양한 개인화 서비스에서 상호운용성을 향상시키기 위해 사용자 선호정보를 의미적으로 공유하는 방법을 제공하는 것이다. 제안하는 접근방법의 가장 큰 특징은 각 서비스의 특성을 반영한 서비스 온톨로지(Service-specific Ontology)들과 이러한 온톨로지들을 기반으로 상호작용할

수 있는 매커니즘을 제공함으로써 사용자의 선호정보를 공유하는 것이다. 즉, 획득한 선호정보를 자신의 서비스에서만 뿐만 아니라 다른 개인화서비스에서도 접근하고 사용 가능하도록 하는 것이다.

본 논문에서 제안하는 시맨틱웹 기반 사용자 선호정보 공유 모델은 그림 2와 같이 3계층 모델을 기반으로 한다. 이러한 3계층 모델은 사용자 선호정보를 관리하는 미들웨어를 통해 사용자 선호정보와 이를 이용하는 사용자 응용 서비스 사이의 독립성과 유연성을 보장한다.

응용계층(Application Tier)에서는 분산 환경의 다양한 개인화서비스들이 각 서비스의 특성에 맞게 서비스 온톨로지를 정의하고, 이를 기준으로 정보나 콘텐츠를 분류하여 제공한다. 제안모델에서는 사용자의 선호정보를 공유하기 위해 데이터계층(Data Tier)에서 사용자의 선호정보를 각 서비스 온톨로지를 기반으로 UPDL(User Preference Description Language)로 기술하고 사용자 프로파일(RDF 문서)에 저장함으로써, 사용자 선호정보를 의미적으로 공유한다. UPDL은 사용자가 자신의 선호정보를 서비스 온톨로지 상에서 기술하는 것을 가능하게 하는 본 논문에서 제안하는 온톨로지 기반의 간단명료한 기술언어(Description Language)이다. 미들웨어계층(Middleware Tier)에 위치하는 UPMS(User Preference Management System)는 사용자와 각 서비스제공자에게 사용자 선호정보를 공유하기 위한 사용자 인터페이스(UI)와 프로그래밍 인터페이스(API)를 제공한다. 제안하는 접근방법을 지원하기 위한 각 요소기술들에 대해서는 4장에서 자세히 설명한다.

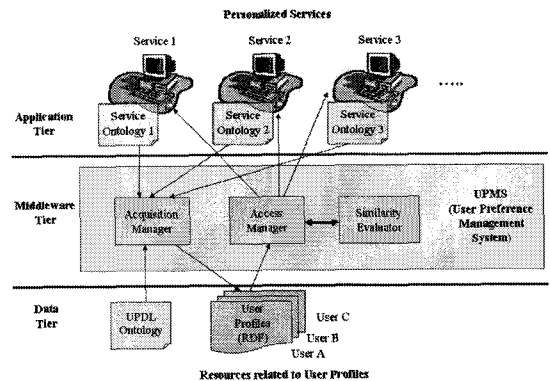


그림 2. 시맨틱웹 기반의 선호정보 공유 모델

제안모델은 기존 접근방법과 비교하여 다음과 같은 두 가지 개선점을 제공하는 것을 목표로 제안되었다. 첫째, 제안하는 접근방법은 서비스 온톨로지들을 기반으로 사용자의 선호정보를 명세함으로써 사용자의 선호정보의 표현에 있어 풍부한 용어와 의미를 제공한다. 즉, 서비스 온톨로지는 각 서비스의 목적과 특성에 따라 분류되어지고 명세화되어지기 때문에 서비스 온톨로지는 사용자의 선호정보를 보다 정확하게 기술할 수 있다. 둘째, 제안하는 접근방법은 다양한 개인화 서비스들 사이에서 강력한 상호운용성을 제공한다. 제안하는 모델에서 사용자의 선호정보는 OWL 기반의 UPDL(User Preference Description Language)과 각 서비스 온톨로지들에 의해 기술되기 때문에, 사용되는 용어에 대한 의미적인 분석이 가능하다. 즉, 사용자 프로파일에 명시된 사용자 관심에 대한 용어가 현재 이용하는 개인화 서비스의 온톨로지상에 정의되어있는 용어와 일치하지 않더라도, 온톨로지상에서 해당 용어의 의미에 대한 유사성 분석을 통해 의미적인 공유가 가능하다.

본 절에서 언급한 바와 같이 제안하는 접근방법을 위해서는 각 서비스들의 정보나 콘텐츠에 대해 기술하기 위한 서비스 온톨로지와 각 서비스 온톨로지를 기반으로 사용자의 선호정보를 기술하기 위한 프로파일링 기법이 필요하다. 또한 이렇게 기술된 사용자 선호정보를 공유하기 위한 미들웨어가 요구된다. 3.2절부터 3.4절까지는 이러한 시맨틱웹 기반의 사용자 선호정보 공유모델을 지원하기 위한 요소기술들에 대해 구체적으로 기술한다.

3.2 서비스 온톨로지

도메인 온톨로지(Domain Ontology)는 각 도메인을 표현하기 위해서 필요한 개념들과 그 관계를 정의하는데, 이러한 도메인 온톨로지는 각 도메인이나 해당 도메인의 표준화 기관에 의해 정의되어질 수 있다. 각 도메인 온톨로지를 정의하기 위해서 개념(concept)을 계층적으로 정의하고 서로 다른 개념과의 관계를 속성으로 정의해 줌으로써 각 도메인에 대한 정보를 좀 더 정확하게 의미적으로 표현할 수 있다. 본 논문에서 사용하는 서비스 온톨로지(Service-specific Ontology)는 일종의 도메인 온톨로지로서, 각 서비스에서 제공하는 정보나 콘텐츠에 대한 분류와 메타데이터를 기술하는 온톨로지로서 정의

할 수 있다.

서비스 온톨로지를 정의하기 위해서 SUMO와 같은 기존의 메타데이터와 표준들이 사용될 수 있는데, 이와 같은 표준들은 다양한 도메인 온톨로지들 사이에서 상위레벨(high-level)에서의 의미적인 연관성을 제공함으로써 다양한 도메인들 사이에 상호운용성을 향상시킬 수 있는 방법을 제공한다. SUMO(Suggested Upper Merged Ontology) 온톨로지는 IEEE Standard Upper Ontology Working Group에 의해 만들어졌으며, 이 워킹그룹의 목적은 데이터 상호운용성, 정보의 검색, 자동화된 추론 및 자연어 처리를 증진시키기 위한 표준 upper ontology를 개발하는 것이다. 한편, MILO(Mid-Level Ontology) 온톨로지는 상위레벨 온톨로지와 하위레벨(low-level)의 상세한 개념들을 정의하는 도메인 온톨로지들 사이에서 가교의 역할을 담당하는 중간레벨(mid-level) 온톨로지이다. 컴퓨터과학 분야의 경우, ACM Computing Classification System(CSS)이 상위레벨 온톨로지로서 사용되어질 수 있는 등 기존의 메타데이터를 활용함으로써 도메인들 사이의 상호운용성을 향상시킬 수 있다.

본 논문에서는 사용자가 이용하는 다양한 서비스들은 각 서비스 온톨로지를 기반으로 정보 및 콘텐츠를 분류하고 이를 기반으로 서비스를 제공하는 것을 가정한다. 일반적으로 서비스 제공자는 기존의 표준이나 메타데이터를 서비스 온톨로지의 상위레벨로 사용하고 각 서비스의 특성에 따라 온톨로지를 상세화하여 서비스 온톨로지를 정의한다. 이런 경우 본 논문의 제안방식은 사용자 선호정보의 공유에 있어서 보다 큰 효과를 기대할 수 있으며, 만약 기존의 표준이나 메타데이터를 고려하지 않더라도 제안모델은 사용자 선호정보의 의미적 공유를 통하여 그 효과를 발휘할 수 있다.

3.3 사용자 선호정보의 표현

사용자가 자신의 선호정보를 다양한 서비스에서 공유하여 사용하기 위해서는 다양한 서비스와 그 서비스에서 제공하는 아이템 항목들을 고려한 관심의 대상을 기술하기 위한 유연한 표현방법이 필요하다. 본 논문에서 제안하는 UPDL(User Preference Description Language)은 사용자가 자신의 선호정보를 서비스 온톨로지들 상에서 표현할 수 있도록

해주는 일종의 간단명료한 언어로서, OWL(Web Ontology Language)를 기반으로 정의되어진다. 제안하는 UPDL의 목적은 사용자가 자신의 선호항목과 그 가중치를 명세함에 있어서 풍부한 표현력을 제공하고, 이렇게 표현된 사용자 선호 정보들의 공유에 있어서 상호운용성을 제공하는 것이다. 이것은 각 서비스가 포함하는 개념과 그들간의 관계를 명시하는 다양한 서비스 온톨로지들을 참조하여 사용자의 선호정보를 표현할 수 있는 방법을 제공함으로써 가능하다.

그림 3은 서비스 온톨로지 상에서 UPDL에 의해 기술된 사용자 프로파일(RDF 문서)의 일부분을 예시한 것으로, 본 절에서는 UPDL 온톨로지의 핵심적인 클래스와 프로퍼티에 대해서 기술한다. UPDL은 사용자의 다양한 선호정보에 대한 리스트와 각 선호정보에 대한 상세 명세들을 분리하여 명세화하고, 이들을 UPDL에서 정의된 프로퍼티로 연관시킨다. 이를 통해 한명의 사용자에 대해 다양한 선호정보에 대한 상세기술이 가능하기 때문에, 선호정보의 표현에 있어 유연성과 확장성을 제공할 수 있다. UPDL 온톨로지의 대표적인 클래스로는 Profile 클래스와 Preference 클래스가 있다. Profile 클래스는 사용자의 개인정보와 선호정보 리스트를 명세화하기 위해 사용되며, Preference 클래스는 각 선호정보에 대한 상세명세를 기술하기 위해 사용된다.

그림 3의 예에서, Profile 클래스는 아래에 설명하는 두 개의 프로퍼티에 의해 사용자의 개인정보 및 선호정보에 대한 상세명세와 연관시킨다. 하지만, 실제 사용자 선호정보에 대한 상세는 Preference 클래스의 인스턴스에 의해 기술되어진다.

- updl:personalInfo. 이 프로퍼티는 Profile 클래스의 인스턴스와 사용자의 개인정보를 연관시키기 위해 사용되어진다. 예에서, 이 프로퍼티는 'UserA' 사

용자와 UserA의 개인정보 URI인 '&person:UserA'와 연관시킨다. 여기서, 사용자의 개인정보는 VCard 온톨로지나 FOAF 메타데이터와 같은 개인정보를 위한 온톨로지를 기반으로 표현할 수 있다.

- updl:hasPreference. 이 프로퍼티는 다양한 사용자의 선호정보들을 나열하고, 각 선호정보에 대한 상세명세를 표현하는 Preference 클래스의 인스턴스와 연관시키기 위해 사용되어진다. 예에서, 이 프로퍼티는 사용자 'UserA'와 두 개의 Preference 클래스의 인스턴스인 'UserA_InterestingField1'과 'UserA_NotInterestingField2'와 연관시킨다.

Preference 클래스는 각 서비스의 연관된 아이템을 가중치와 함께 사용자의 관심 정보를 기술하기 위해 사용되어진다. 실제로, 사용자 선호정보는 Preference 클래스의 두 개의 하위 클래스인 Interesting 클래스, NotInteresting 클래스에 의해 기술되어지는데, 각각 관심이 있거나 관심이 없는 아이템과 그 가중치를 표현하기 위한 클래스이다. Preference 클래스는 아래의 두 개의 프로퍼티를 사용하여 사용자의 선호정보를 기술한다.

- updl:hasItem. 이 프로퍼티는 사용자의 관심항목과 서비스 온톨로지에 계층적으로 분류되어진 클래스를 연관시키기 위해 사용되어진다. 이 프로퍼티는 URI를 사용하여 관심항목을 연관시키기 때문에, 다양한 서비스 온톨로지들에서 정의된 클래스들을 함께 표현할 수 있다. 예에서, 이 프로퍼티는 하나의 서비스에 대한 컨텐츠를 계층적으로 분류한 'service1' 온톨로지(&service1)의 'DigitalCamera'라는 클래스와 연관시킨다.

- updl:hasWeight. 이 프로퍼티는 관심항목에 대한 가중치를 표현하기 위해 사용되어진다. 본 논문에서는 처음 등록된 사용자의 관심항목에는 가중치 '10' 또는 '-10'을 기본값으로 부여한다. 예에서는 '&service1:DigitalCamera' 관심항목에 대한 가중치가 10이라는 것을 보여준다.

본 절에서 설명한 것처럼, 하나의 특정한 서비스에 대한 선호정보를 표현하는 기존의 연구와 차별되는 UPDL의 가장 큰 특징은 서로 다른 목적과 형태를 가지는 다양한 서비스에 대한 온톨로지들과 그 온톨로지에 대한 URI를 참조하여 사용자의 선호정보를 기술할 수 있도록 하는 것이다. 이러한 접근방법은 사용자의 선호정보를 표현함에 있어 풍부한 표현력

```

<updl:Profile rdf:ID="UserA">
  <updl:personalInfo rdf:resource="#&person:UserA">
  <updl:hasPreference rdf:resource="#UserA_InterestingItem1">
  <updl:hasPreference rdf:resource="#UserA_NotInterestingItem1">
</updl:Profile>

<updl:Interesting rdf:ID="UserA_InterestingItem1">
  <updl:hasItem rdf:resource="#&service1:DigitalCamera">
  <updl:hasWeight> 10 </updl:hasWeight>
</updl:Interesting>
    
```

그림 3. 제안모델에서의 개인선호정보 표현

을 제공할 수 있도록 하며, 다양한 서비스들간에 사용자 선호정보를 공유할 수 있도록 한다. 따라서 UPDL은 어떠한 서비스에서 사용되는 사용자의 선호정보도 동적으로 표현하고 공유할 수 있다.

3.4 사용자 선호정보의 공유

본 절에서는 3.3절에서 설명한 사용자 프로파일링 기반으로 사용자의 선호정보 공유를 지원하기 위한 미들웨어인 UPMS(User Preference Management System)에 대해 기술한다. UPMS는 Acquisition Manger, Access Manager와 Similarity Evaluator를 통하여 사용자/프로그래밍 인터페이스를 제공한다. 사용자나 다양한 개인화서비스들은 UPMS를 통하여 사용자 프로파일에 대한 직접적인 접근이나 처리 없이 사용자의 선호정보를 입력하고 접근할 수 있으며, 다른 다양한 서비스들의 사용자의 선호정보를 공유할 수 있다(3장 그림 2 참조). Acquisition Manager는 사용자가 서비스 온톨로지 상에서 자신의 선호정보를 UPDL에 따라 사용자 프로파일에 입력하기 위한 인터페이스를 제공한다. Acquisition Manager는 서비스 온톨로지의 브라우저를 통해 사용자로부터 선호정보 획득(그림 4의 좌)하는 과정과 획득된 정보를 UPDL에 의해 기술하고 사용자 프로파일에 RDF형식으로 자동으로 갱신(그림 4의 우)하는 과정으로 이루어진다.

Access Manager는 다양한 개인화 서비스들이 RDF로 기술된 사용자 프로파일로부터 개인화 서비스를 위해 필요한 사용자의 선호정보를 이용할 수 있는 프로그래밍 인터페이스(API)를 제공한다. Similarity Evaluator는 Access Manager에 의해 사용되는데, 사용자 프로파일에 등록된 사용자의 선호정보의 관심항목에 대한 현재 요청중인 서비스 온톨로지 상에서의 유사성을 기반으로, 각 선호정보에 대한 요청 서비스에서의 의미적으로 가장 유사한 분류와 그 가중치를 평가한다. 이러한 사용자 선호정보의 유사성 평가를 위해서 다양한 방법들이 사용되어질 수 있지만, 본 논문에서는 온톨로지의 특성을 반영하여 동일한 클래스의 부재 시에 그림 5와 같이 상위 클래스를 반복 비교하면서 상위클래스의 평가가중치는 하위 클래스 평가가중치의 50%로 계산한다. 3.2절에서 기술한 서비스 온톨로지의 특성에 따라 상위 클래스는 하위 클래스보다 일반화된 의미를 가지

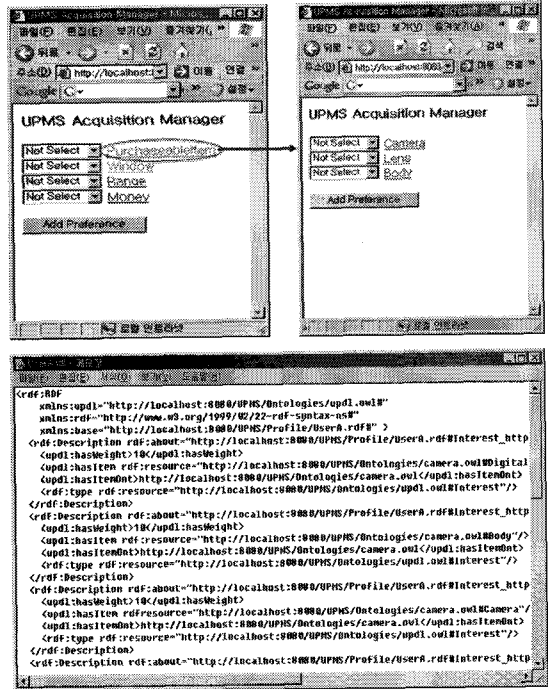


그림 4. UPMS Acquisition Manager의 인터페이스(상)와 생성된 사용자 프로파일의 예(하)

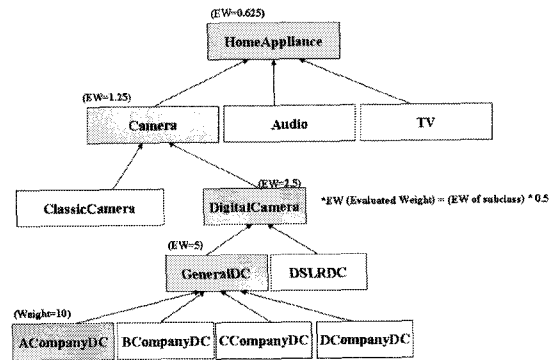


그림 5. 평가가중치 계산

며, 하위 클래스는 상위 클래스보다 구체적인 의미를 가지게 된다. 따라서 평가된 유사클래스가 등록된 사용자의 관심항목보다 일반화된 의미를 가지는 상위 클래스일 경우 사용자의 구체적인 실제 선호와는 차이가 있을 가능성은 높아지기 때문에, 본 논문에서는 그 평가가중치를 사용자의 관심항목과의 거리에 따라 50%씩 점진적으로 감소시킨다.

UPMS 미들웨어에서 온톨로지와 RDF문서의 처리를 위해 Jena Toolkit [18]을 사용하여 구현하였다.

Jena Toolkit은 OWL과 같은 온톨로지와 RDF 문서를 생성하는 Java 응용 프로그래밍 인터페이스(API)를 제공한다.

4. 제안 시스템의 분석 및 평가

4장에서는 제안 시스템에 대한 분석과 평가를 기술한다. 본 논문에서 제안한 시맨틱 웹 기반의 사용자 선호정보 공유 모델은 개인화서비스를 제공하기 위해 요구되는 사용자 선호정보를 다양한 서비스들이 의미적으로 공유할 수 있도록 해준다. 예를 들어, 전자제품을 판매하는 전자상거래 서비스에서 사용자가 A사 디지털 카메라를 관심항목에 등록한 경우, 인터넷 서점에서 이 정보를 공유하여 디지털 카메라 관련서적을 추천해 줄 수 있다. 제안모델은 기존 연구에 비해 다음과 같은 장점을 제공한다.

- 상호운용성 제공: 제안모델은 사용자의 선호정보를 서비스 온톨로지 상에서 OWL기반 UPDL에 의해 기술하기 때문에, 다양한 개인화서비스들 사이에서 사용자 선호정보의 공유를 지원한다.
- 풍부한 표현 가능: 제안모델은 사용자의 선호정보를 각 서비스 온톨로지 상에서 기술하는 것이 가능하기 때문에, 사용자의 관심항목을 의미를 기반으로 보다 풍부한 용어로서 기술할 수 있다. 또한, 서비스 온톨로지는 UPMS에서 필요시 실시간으로 참조하기 때문에, 서비스 온톨로지의 변화도 유연하게 반영할 수 있다.
- 개인화서비스 품질의 향상: 제안 모델은 계층적으로 분류된 서비스 온톨로지를 기반으로 사용자 선호정보를 기술하기 때문에, 완전히 일치하는 항목뿐만 아니라 유사한 의미를 가지는 항목에 대해서도 추천이 가능하다.
- 유연성 제공: 새로운 서비스가 추가되거나 기존 참여 서비스의 온톨로지가 변경되더라도, 기존의 서비스나 UPMS의 변화 없이 이를 즉시 반영할 수 있다.

제안 시스템을 평가하기 위해서 5개의 서비스(인터넷쇼핑몰, 인터넷서점, 전자제품 상품소개 서비스, 카메라 정보제공 서비스, 디지털카메라온라인강의 서비스)에 대해서 테스트 온톨로지를 간단하게 정의하고, 이를 기반으로 테스트를 위한 각 서비스를 시뮬레이션 하였다. 구현된 각 서비스에서 UPMS의

Aquisition Manger를 이용하여 서비스 온톨로지 상에서 각 서비스 당 2개씩의 선호정보를 등록하여 모두 10개의 선호정보를 등록하고 테스트를 수행하였다. 성능평가에서 명확한 결과를 얻기 위해, 등록하는 각 선호정보는 서로 정확하게 일치하지는 않는 항목들로 선택하였다. 그림 6은 테스트를 위해 사용된 서비스와 서비스 온톨로지 및 사용자 프로파일에 등록된 선호정보들의 평가가중치 결과의 예를 보여주고 있다.

그림 6에서 제안시스템은 다음과 같은 과정을 통해 사용자 선호정보를 공유하고 활용되어진다. (1) 사용자는 각 서비스에서 자신의 선호정보를 등록하기 위해 해당 서비스의 선택메뉴를 통해 UPMS의 Acquisition Manager로 접근하게 되고, UPMS는 3.4절의 그림 4와 같이 현재 서비스 온톨로지 기반의 브라우저를 제공하는 사용자 인터페이스를 통해 사용자의 선호정보를 획득하여 해당 사용자의 프로파일에 등록한다. (2) 사용자는 다른 서비스에서 자신이 관심있는 정보를 추천받기 위해 그림 6과 같이 UPMS 추천메뉴를 통해 UPMS의 Access Manager에 접근한다. 이때 UPMS는 사용자 선호정보의 서비스 온톨로지와 현재 요청한 서비스의 서비스 온톨로지 상에서 3.4절의 그림 5와 같은 Similarity Evaluator의 유사성 평가를 통해 평가항목과 그 평가가중치를 요청서비스에 전달한다. 그림 6의 UPMS 처리결과를 사용자 프로파일에 저장된 선호항목/가중치와 Similarity Evaluator에 의해 평가된 평가항목/평가가중치를 비교하여 출력한 화면이다.

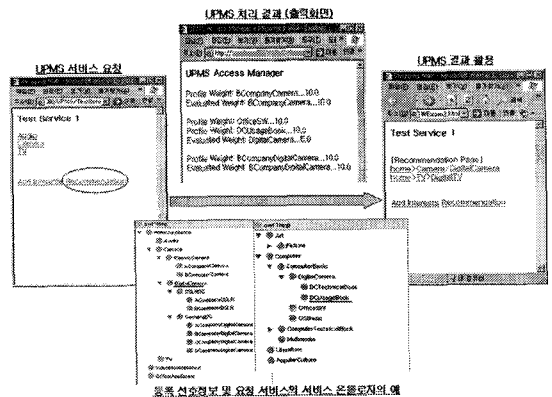


그림 6. 테스트 서비스의 예 (좌, 우)와 온톨로지 (하) 및 UPMS Access Manager의 선호정보와 평가가중치(상)

(3) 요청서비스는 전달받은 평가항목/평가가중치를 이용하여 사용자가 선호할 가능성이 높은 콘텐츠를 추천한다. 이때 다양한 추천 알고리즘이 사용될수 있는데, 그림 6은 평가가중치에 대한 랭킹으로 추천한 결과이다.

표 1은 참여서비스가 5개인 경우의 각 평가조건에 해당하는 개인화서비스를 제공하기 위해 가용한 선호정보의 개수를 나타낸 것이다. 표 1에서 보는 것처럼 기존방식의 경우 단일시스템 내에서 사용자의 선호정보를 관리하고 사용하기 때문에 참여서비스의 수와 상관없이 동일한 결과를 나타낸다. 즉, 개인화서비스를 위해 자신의 서비스에서 등록된 사용자 선호정보만을 사용할 수 있다. 반면에, 제안모델은 어떠한 서비스 사이에서도 사용자의 선호정보를 공유할 수 있는 기반을 제공하기 때문에, 모두 서비스에서 10개의 선호정보를 개인화서비스를 위해 사용 가능하다. 하지만, 서로 다른 서비스들에서 등록된 사용자 선호정보라 하더라도 의미적으로 정확히 일치하는 항목이 존재할 수 있기 때문에, 실제의 경우에는 2개에서 10개 사이의 가용한 선호정보가 존재하게 될 것이다.

등록된 선호정보가 가용한 것이라 하더라도, 실제로는 평가가중치가 일정 값 이상의 값을 가지는 경우에만 유사성이 존재하므로 개인화서비스에서 사용자에게 제공하는 결과에 영향을 주게 된다. 표 1에서 평가가중치 1이상인 가용한 선호정보 개수의 평균은 7.2개, 평가가중치 2이상인 가용한 선호정보 개수의 평균은 5.6개, 평가가중치 5이상인 가용한 선호정보의 개수는 평균 3.2개로 모두 자신의 서비스에서 등록된 선호정보(2개) 이외에 타 서비스에서 등록된 선호정보도 자신의 개인화서비스를 위해 사용될 수 있음을 알 수 있다.

그림 7은 참여서비스 수 증가에 따라 평균 가용한 선호정보의 수를 나타낸 그래프로써, 참여서비스가

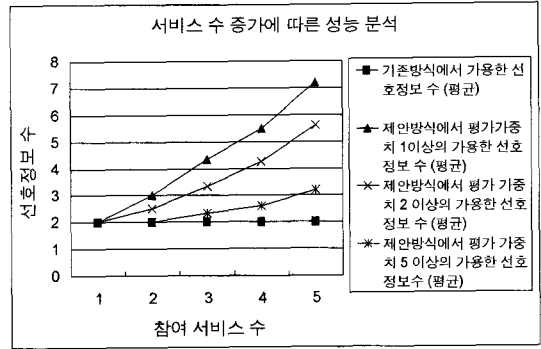


그림 7. 가용한 선호정보 수의 변화

증가함에 따라 가용한 선호정보의 수가 증가함을 볼 수 있다. 결과적으로, 참여하는 서비스들 서로간의 연관성에 따라 정도의 차이는 있겠지만, 참여서비스의 수가 증가할수록 제안모델의 유용성은 크게 증가할 것으로 예상된다.

5. 결 론

급증하는 정보와 다양한 인터넷 서비스의 홍수 속에서 개인화 서비스(Personalized Services)의 중요성과 함께 다양한 서비스들 사이에서의 사용자 선호정보의 공유는 개인화서비스의 품질 향상의 중요한 요소가 될 것이다. 본 논문에서는 사용자의 선호정보를 공유하기 위한 방법으로 시맨틱웹 기반의 사용자 선호정보 공유 모델을 제안했다. 제안모델은 서비스 온톨로지들을 기반으로 사용자의 선호정보를 기술(describe)할 수 있도록 UPDL(User Preference Description Language)를 정의하고, 사용자의 선호정보를 의미기반의 풍부한 용어와 함께 정확하게 기술할 수 있는 방법을 제공한다. 또한 미들웨어인 UPMS(User Preference Management System)를 통해 다양한 개인화 서비스들 사이에서 사용자 선호

표 1. 가용한 선호정보 수 비교 (참여서비스: 5개)

| 평가항목 | 서비스1 | 서비스2 | 서비스3 | 서비스4 | 서비스5 | 평균 |
|-------------------|------|------|------|------|------|-----|
| 단일시스템 | 2 | 2 | 2 | 2 | 2 | 2 |
| 제안방식의 평가가중치(1이상) | 8 | 7 | 8 | 6 | 7 | 7.2 |
| 제안방식의 평가가중치(2 이상) | 6 | 6 | 7 | 4 | 5 | 5.6 |
| 제안방식의 평가가중치(5 이상) | 3 | 4 | 4 | 2 | 3 | 3.2 |

정보의 의미적 분석을 통해 강력한 상호운용성(interoperability)을 제공한다.

본 논문의 제안모델은 다음과 같은 장점을 제공한다. 첫째, 제안모델은 다양한 개인화서비스들 사이에 사용자 선호정보의 의미적인 공유를 지원한다. 둘째, 제안모델은 사용자의 관심항목을 각 서비스의 특성을 반영하여 의미기반의 보다 풍부한 용어로 기술할 수 있다. 셋째, 제안모델은 완전히 일치하는 항목뿐 아니라 유사한 항목에 대해서도 추천이 가능하다. 넷째, 제안모델은 서비스의 추가나 변경에도 유연한 적응력을 가진다.

제안모델은 현재의 인터넷 환경에서 사용자에게 편리한 서비스 이용을 가능하게 할 뿐 아니라 제한된 성능, 제한된 사용자 인터페이스 및 낮은 대역폭의 통신환경을 제공하는 단말기에서 분산된 다양한 서비스에 대한 접근이 많이 이루어질 것으로 예상되는 유비쿼터스 환경에서 더욱 유용할 것으로 사료된다. 향후 연구로는 다양한 분야의 개인화서비스에서 제안모델의 적용과 실제 환경에서의 가용성에 대한 평가가 필요할 것으로 본다.

참 고 문 헌

- [1] Maurice D. Mulvenna, Sarabjot S. Anand, and Alex G. Bochner, "Personalization on the Net using Web Mining: Introduction," *ACM Communication*, Vol.43, No.8, pp. 122-125, 2000.
- [2] D. Riecken, "Personalized Views of Personalization," *ACM Communication*, Vol.43, No.8 pp. 27-28, 2000.
- [3] T. Berners-Lee, et.al., *The Semantic Web*, Scientific American, 2001.
- [4] Antoniou and van Harmelen. *A Semantic web Primer*, The MIT Press, Cambridge, 2004.
- [5] Dwi H. Widyantoro, Thomas R. Ioerger, and John Yen, "Learning User Interest Dynamics with a Three-Descriptor Representation", *Journal of the American Society of Information Science*, Vol.52, No.3, pp. 212-225, 2001.
- [6] S.K.M. Wong and C.J. Butz, "A Bayesian Approach to User Profiling in Information Retrieval," *Technology Letters*, Vol.4, No.1, pp. 50-56, 2000.
- [7] Donghai Guan, Qing Li, Sungyoung Lee, and Youngkoo Lee, "A Context-Aware Music Recommendation Agent in Smart Office," *Lecture Notes in Artificial Intelligence*, Vol.4223, pp. 1201-1204, 2006.
- [8] Steffen Staab and Hannes Werthner, "Intelligent Systems for Tourism," *IEEE Intelligent Systems*, Vol.17, No.6, pp. 53-66, 2002.
- [9] Shijun Yu, et.al., "User Profiles in Location-based Services: Make Humans More Nomadic and Personalized," *Databases and Applications*, ACTA Press, pp. 24-30, 2004.
- [10] Mike Collett(Chair), *Draft Standard for Learning Technology - Public and Private Information (PAPI) for Learners (PAPI Learner)*, IEEE, New York, USA, 2000.
- [11] IMS Global Learning Consortium, *IMS Learner Information Package Specification*, IMS, 2003.
- [12] Julien Tane, Christoph Schmitz, and Gerd Stumme, "Semantic Resource Management for the Web: An E-Learning Application," *WWW2004*, *ACM Press*, New York, NY, USA, pp. 1-10, 2004.
- [13] Yugyung Lee, Chintan Patel, Soon Ae Chun, and James Geller., "Compositional Knowledge Management for Medical Services on Semantic Web," *WWW2004*. *ACM Press*, New York, NY, USA, pp. 498-499, 2004.
- [14] Jong-Woo Kim, Ju-Yeon Kim, Hyun-Suk Hwang, and Chang-Soo Kim, "Location-Sensitive Tour Guide Services using the Semantic Web," *Lecture Notes in Computer Science*, Vol.3682, pp. 908-914, 2005.
- [15] Frank Manola and Eric Miller, *RDF Primer* (<http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>), W3C (MIT, ERCIM, Keio), 2004.
- [16] Stefan Decler, Prasenjit Mitra, and Sergey Melnik, "Framework for the Semantic web:

An RDF Tutorial," *IEEE Internet Computing*, Vol.4, No.6, pp. 68-73, 2000.

- [17] Michael K. Smith, Chris Welty, and Deborah L. McGuinness, *OWL Web Ontology Language Guide* (<http://www.w3.org/TR/owl-guide/>), W3C (MIT, ERCIM, Keio), 2004.
- [18] Jeremy J. Carroll, Ian Dickinson, and Chris Dollin, "Jena: Implementing the Semantic Web Recommendation," *WWW2004, ACM*, New York, USA, pp. 74-83, 2004.



김 주 연

2000년 한국방송대학교 전자계산학과 이학사
 2003년 경성대학교 정보공학과 공학석사
 2006년~현재 부경대학교 정보보호학(협) 박사과정
 관심분야 : 개인화서비스, 시맨틱 웹



김 종 우

1998년 부경대학교 전자계산학과 이학사
 2003년 부경대학교 전자계산학과 이학석사
 2006년 부경대학교 정보보호학(협) 공학박사
 2007년~현재 대구대학교 정보통신공학부 누리초빙교수

관심분야 : 시맨틱 웹, LBS, GIS



김 창 수

1984년 울산대학교 전자계산학과 학사
 1986년 중앙대학교 컴퓨터공학과 석사
 1991년 중앙대학교 컴퓨터공학과 박사
 1992년~현재 부경대학교 전자컴퓨터 정보통신공학부 교수

관심분야 : 운영체제, 임베디드 시스템, 이동통신, LBS/GIS, 시맨틱 웹, 도시방재시스템 등