

시맨틱 기능을 활용한 oneM2M 기반 스마트시티 링크드 오픈 데이터

이지은 세종대학교 정보보호학 박사, 전자부품연구원 자율지능이연구센터 연구원
송재승 서울인터넷특별기술위원회(STC 1) 의장, 세종대학교 정보보호학 교수

1. 머리말

사물인터넷 기반의 스마트시티는 도시에 설치된 각종 센서에서 수집한 데이터를 활용하여 다양한 서비스를 제공하고 시민의 삶의 질을 높여주는 도시를 말한다. 하지만 스마트시티의 단순 사물인터넷 플랫폼을 활용해서 수집된 원시 데이터들은 데이터의 의미나 다른 데이터와의 관계를 명시하는 메타데이터가 없다. 이 때문에 서로 다른 산업이나 서비스 간의 데이터 공유와 새로운 서비스 창출을 위한 데이터 재활용에 어려움이 있었다. 링크드 오픈 데이터는 URI를 통해 모든 데이터의 연결을 가능하게 한다. 그리하여 스마트시티 공공 데이터의 통합과 개방으로 도시 데이터가 재사용될 수 있게 한다. 데이터 간의 연결을 가능하게 하는 시맨틱 데이터는 이러한 링크드 오픈 데이터의 기반이 되어 스마트시티 내에 공존하는 여러 도메인 간의 상호운용성과 데이터 이질성 문제를 해결한다. 이에 따라 oneM2M 사물인터넷 국제표준화 기구에서도 다양한 스마트시티 도메인 온톨로지를 개발하고 있다. 또한 시맨틱 데이터 모델을 정의하는 등 스마트시티의 시맨틱 기술 적용을 위한 표준화 활동에 힘쓰고 있다. 본고에서 논의하는 스마

트시티 링크드 오픈 데이터 시스템은 oneM2M의 시맨틱 기능을 활용하여 공공 데이터의 통합과 연결성을 지원하고, 이를 개방하여 시민들이 사용할 수 있게 함으로써 도시 데이터의 높은 활용도를 달성한다. 또한 실시간 데이터 연결성의 시각적 표현과 검색 기능 제공을 통해 스마트시티 내에서 일어나고 있는 도시 문제를 쉽게 발견할 수 있다.

2. 시맨틱과 링크드 오픈 데이터 및 oneM2M 시맨틱스

스마트시티 링크드 오픈 데이터(LOD, Linked Open Data)는 스마트시티의 사물인터넷 플랫폼에서 수집되는 다양한 데이터를 의미를 통해 연결하고, 공유가 가능하도록 한다. 이로써 데이터의 재사용 및 새로운 가치를 제공하는 데이터를 창출할 수 있도록 개방형 통합 데이터를 제공하는 서비스이다. 스마트시티 LOD 시스템은 사물인터넷 국제표준인 oneM2M에서 제공하는 시맨틱 주석화 등의 기능을 포함하는 시맨틱 모듈이 도시에서 발생하는 수많은 공공 데이터를 사물인터넷 플랫폼의 시맨틱 데이터와 의미적으로 연결한다. 그럼으로써 서로 다른 도메

인들의 연결 및 새로운 서비스 창출이 가능하다. 다음 단락에서는 스마트시티 LOD 시스템의 기반이 되는 기술들에 대해 서술한다.

2.1 시맨틱스(Semantics)

시맨틱 웹은 World Wide Web Consortium (W3C)의 창시자인 팀 버너스리가 처음 제안한 개념으로[1], RDF(Resource Description Framework), OWL(Web Ontology Language)과 같은 공통 어휘를 사용함으로써 데이터의 의미적 정보를 기계가 처리할 수 있게 하는 기술이다. 시맨틱스는 사람이나 사물과 같은 데이터 자원 사이의 관계 형성을 가능하게 하고 URI(Uniform Resource Identifier)를 통해 모든 데이터가 연결되도록 한다. 다음 단락에서는 스마트시티 LOD 시스템 시맨틱 데이터의 획득과 표현 기술에 대해 서술한다.

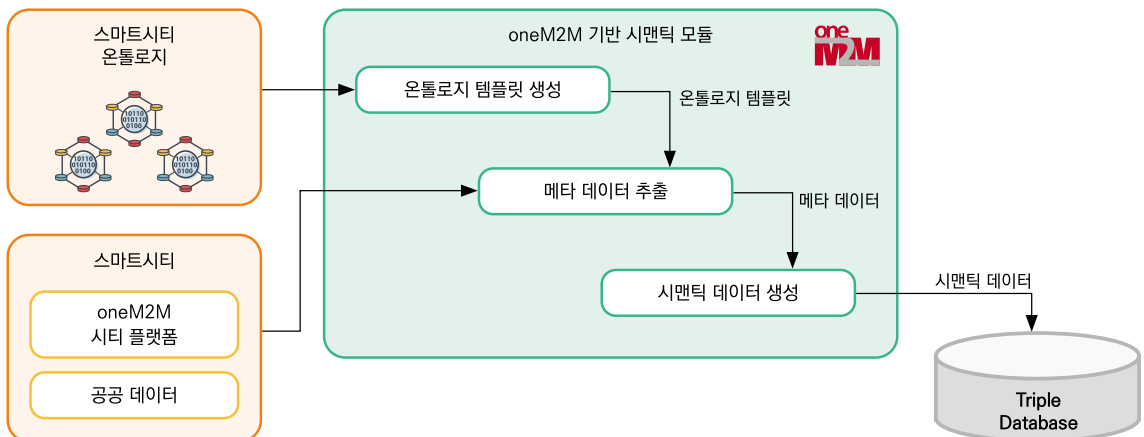
2.1.1 시맨틱 데이터 획득

스마트시티 LOD 시스템의 시맨틱 데이터 획득(Semantic Data Acquisition)은 스마트시티의 사물 인터넷 플랫폼 및 데이터베이스 역할을 하는 데이터 허브에서 공공 데이터를 수집하여 시맨틱 데이터를 생성하는 것을 말한다. [그림 1]과 같이 oneM2M

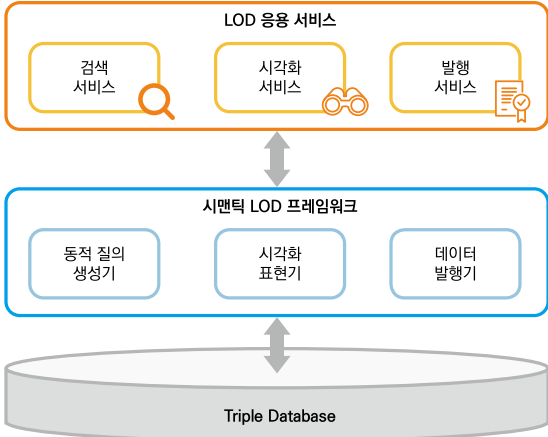
표준 기반의 스마트시티 플랫폼에서 수집되는 공공 데이터에서 메타 데이터를 추출한다. 다음으로 스마트시티 온톨로지(Ontology)를 기반으로 추출한 메타 데이터를 온톨로지 템플릿에 따라 매핑함으로써 oneM2M 기반 시맨틱 모듈이 시맨틱 데이터를 획득하는 역할을 한다. oneM2M 기반 시맨틱 모듈은 도시 데이터를 실시간 시맨틱 데이터로 구축하여 시맨틱(혹은 Triple) 데이터 전용 데이터베이스인 TDB(Triple Database)에 시맨틱 데이터를 구축함으로써 데이터 간의 연결을 식별할 수 있게 한다.

2.1.2 시맨틱 데이터 표현

스마트시티의 공공 데이터에서 시맨틱 정보를 추출하기 위해서는 RDF, OWL과 같은 공통 어휘로 작성된 온톨로지가 필요하다. 스마트시티 LOD 시스템의 시맨틱 모듈은 스마트시티에 정의되어 있는 공통 온톨로지를 입력 받아 이를 기반으로 메타 데이터 작성에 필요한 온톨로지 템플릿을 추출한다. 메타 데이터는 스마트시티 서비스의 기반이 되는 엔티티들이 반정형 기반(semi-structured-based) 표현방식으로 구성되어 있으며, 시맨틱 모듈을 통해 엔티티 사이의 관계를 가지는(예: hasRelationship, hasProperty, hasValue 등) 시맨틱 데이터를 표현할 수 있다.



[그림 1] 시맨틱 데이터 획득 과정



[그림 2] 스마트시티 LOD 시스템 아키텍처

2.2 링크드 오픈 데이터

링크드 오픈 데이터는 공공 데이터의 개방과 유통을 위해 RDF 형식의 데이터 표현, SPARQL(Sparkle Protocol and RDF Query Language) 등을 지원함으로써 모든 데이터가 URI를 통해 웹상에서 연결되는 기술이다[2]. 스마트시티 LOD 시스템은 이러한 링크드 오픈 데이터 기술을 위해 스마트시티의 공공 데이터를 메타 데이터로 시맨틱 모델링하여 변환하는 시맨틱 모듈을 활용한다. 그리하여 이(異)기종 데이터를 통합하고 개방하는 것뿐만 아니라 실시간 시맨틱 데이터 검색, 이기종 도메인 간의 연결 관계 시각화로 스마트시티 내에서 일어나고 있는 다양한 도시 문제 발견에 도움을 줄 수 있다.

2.3 oneM2M 시맨틱스

oneM2M에서는 다양한 표준을 통합하고, 데이터에 의미를 추가함으로써 이기종 데이터 간의 호환성을 보장하는 지능형 사물인터넷 스마트시티 서비스를 제시한다. 특히 사물인터넷 데이터의 리소스 구조에 데이터 간의 관계, 의미 등의 메타데이터를 추가할 수 있도록 기준이 되는 기본 온톨로지(Base Ontology) 및 다양한 스마트시티 도메인 온톨로지

를 개발하여 스마트시티에 시맨틱 기술을 적용하였다. 특히, 정형화된 형식의 데이터에 의미 정보를 추가하여 시맨틱 데이터로 변환하는 시맨틱 주석화(Semantic Annotation), 시맨틱 및 컨텍스트 기반 오픈 API를 통한 시맨틱 검색 기능처럼 스마트시티 LOD 시스템에 활용할 수 있는 여러 시맨틱 표준 기술들을 제시하고 있다.

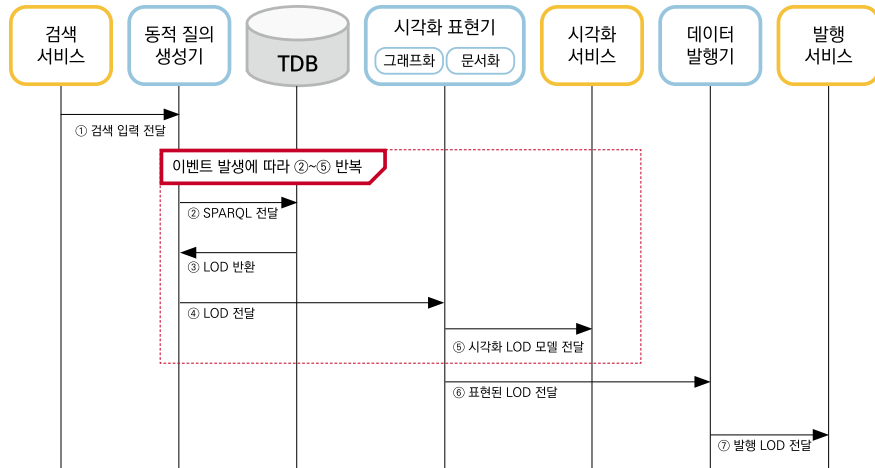
다음은 oneM2M에서 제공하고 있는 표준 기반 시맨틱 기능의 일부를 보여준다.

- **시맨틱 주석화(Semantic Annotation):** oneM2M IoT플랫폼의 자원에 의미정보(메타데이터)를 추가한 <SemanticDescriptor>를 정의함으로써 이기종 IoT 애플리케이션에 의해 주석자원이 의미적으로 발견될 수 있도록 함
- **oneM2M 기본 온톨로지(oneM2M Base Ontology):** oneM2M에서 정의한 온톨로지로서, 다른 IoT 시스템의 다양한 외부 온톨로지를 oneM2M 기본 온톨로지에 매핑하여 oneM2M 시스템과 외부 시스템 간의 상호작용을 지원함
- **시맨틱 매시업(Semantic Mashup):** oneM2M 기본 온톨로지와 다른 외부 온톨로지를 결합하여 새로운 연결 관계 도출을 지원함
- **시맨틱 검증(Semantic Validation):** 온톨로지를 참조하여 oneM2M 표준에 따라 시맨틱 데이터의 유효성 검증을 지원함

참고로 Release 5에서는 시맨틱 기능을 고도화시켜 더 지능화된 시맨틱 검색이 가능하도록 하는 신규 Work Item(WI-0101, Advanced Semantic Discovery)이 제안되어 Telecom Italia, KETI, Inria, EGM, Deutsche Telecom, Orange 등의 주도로 표준화가 진행되고 있다. 특히, 해당 Work Item은 ETSI 산하에 oneM2M 시맨틱 기능에 대한 강화를 목적으로 STF(Special Task Force)를 별도로 구성하고, 여기서 만들어진 STF589의 시맨틱 전문가들이 관련 연구를 수행하면서 도출된 결과물들을 oneM2M에 기고하는 방식으로 진행이 되고 있다.

3. 스마트시티 LOD 시스템

기존의 스마트시티에서는 수많은 센서 데이터가 다양한 도메인 및 서비스에서 생성된다. 따라서



[그림 3] 스마트시티 LOD 시스템 프로시저

동일한 표준을 사용하지 않는 한 서로 다른 도메인 및 서비스 간 이질성으로 인해 상호 운용성이 부족해지며, 이는 도시 데이터의 접근과 재사용성을 제한한다. 본고에서 제안하는 스마트시티 LOD 시스템은 oneM2M에서 제공하는 시맨틱 기능과 기존 LOD 개념을 접목하여 파편화된 스마트시티 데이터들을 연결하고 재사용이 가능하도록 함으로써, 누구나 쉽게 스마트시티 데이터를 검색하고 활용할 수 있도록 지원한다.

스마트시티 LOD 시스템은 [그림 2]와 같이 스마트시티에서 생산되는 시맨틱 데이터를 실시간으로 저장하고 관리하는 시맨틱 기반 데이터베이스인 TDB, 스마트시티 데이터의 연결성을 시맨틱에 대한 전문지식이 없는 일반 사용자도 사용할 수 있도록 관리하는 시맨틱 LOD 프레임워크, LOD 개방 및 공유를 위한 서비스 체계 구축이 가능한 LOD 응용 서비스로 구성되어 있다.

본 시스템은 oneM2M의 시맨틱 주석화 기능을 통해 변환한 시맨틱 데이터를 기초로 하며, 시맨틱 데이터들은 실시간으로 업데이트되어 TDB에 구축된다. 구축된 시맨틱 데이터는 고유한 URI를 통해 모든 데이터가 연결되는 시맨틱 LOD의 기반이 된다. [그림 3]은 스마트시티 LOD 시스템의 응용 서비스 제공

과정이다. 시맨틱 LOD 프레임워크(파랑)와 LOD 응용 서비스(노랑)의 구성요소들이 사용자와 상호작용하여 서비스를 제공한다. LOD 응용 서비스가 사용자로부터 검색을 입력받으면 시맨틱 LOD 프레임워크에서 시각화 표현 정보를 생성하여 서비스한다. 또한 사용자의 입력에 따른 이벤트 발생을 통해 즉각적인 상호작용을 지원한다. 다음 단락에서는 스마트시티 LOD 시스템의 서비스 제공을 위해 시맨틱 LOD 프레임워크와 LOD 응용 서비스의 각 요소들이 수행하는 동작에 대해 자세히 기술한다.

3.1 시맨틱 LOD 프레임워크

시맨틱 LOD 프레임워크는 TDB에 구축되어 있는 시맨틱 LOD에 대한 접근을 가능하게 하며 누구나 LOD 데이터를 이해, 활용할 수 있도록 표현하고 발행해 주는 역할을 한다. 본 프레임워크는 동적 질의의 생성기, 시각화 표현기, 데이터 발행기로 구성되어 있다. 동적 질의의 생성기는 LOD 응용 서비스에서 원하는 검색을 입력 시 정보를 전달받아 동적으로 질의(SPARQL)를 생성하여 TDB에 전달하여 시맨틱 LOD 검색 기능을 자동화한다. 동적 질의의 생성기를 통해 실시간으로

정확하게 검색할 수 있으며, 이기종 도메인 간의 관계성(데이터의 연결성) 검색이 가능하다.

동적 질의 생성기가 생성한 SPARQL 쿼리에 따라 TDB가 시맨틱 LOD를 반환하면 시각화 표현기로 전달한다. 시각화 표현기는 시맨틱 LOD의 의미와 연결 관계를 한눈에 이해할 수 있도록 서로 유기적으로 연결되어 있는 데이터 모델을 그래프와 표의 형태로 시각화함으로써 시맨틱 LOD 데이터의 표현력을 높인다. 특히, 그래프 표현 방식은 입력한 데이터의 검색 결과 이외에도 사용자가 생성하는 이벤트 정보를 동적 질의 생성기에 전달하여 관계성을 드러내 더 깊은 연결성을 선택적으로 표현할 수 있다. 예를 들어 ‘야탑 공영주차장’을 검색했다면 지리적 위치, 잔여 주차 공간과 같은 야탑 공영주차장과 연결된 1레벨 데이터들과의 연결성을 표현하지만, 지리적 위치를 클릭하면 지리적 위치와 연결된 2레벨 데이터들과의 연결성 또한 같이 드러내 연결성을 더 깊이 표현한다.

마지막으로 LOD 발행기는 스마트시티 공공 데이터의 재사용성과 활용성을 높이기 위해 시각화 표현기에서 전달받은 LOD를 통해 데이터 연결 관계와 데이터의 의미를 공유하는 LOD 배포 기능을 제공한다. LOD는 RDF, OWL, N-3 등 여러 가지 형태로 발행될 수 있으며 사용자가 원하는 형태의 데이터 파일을 생성함으로써 스마트시티 LOD 시스템에서 제공하는 응용 서비스 외의 다양한 서비스를 창출한다.

3.2 LOD 응용 서비스

LOD 응용 서비스는 스마트시티 LOD 데이터의 표현과 공유, 검색 서비스를 실제로 제공하는 역할을 함으로써 시맨틱 LOD의 의미 관계 이해와 실시간 스마트시티 도시문제 예측을 가능하게 한다. 본 시스템에서 제안하는 응용 서비스는 크게 세 부분, 즉 LOD 검색, LOD 시각화, LOD 발행으로 구성되어 있다.


LOD 검색 서비스는 사용자가 스마트시티의 LOD

공공 데이터를 검색할 수 있는 서비스이다. 시맨틱이나 LOD에 대한 지식이 없더라도 손쉽게 원하는 키워드 또는 토픽들을 검색하여 이와 연결된 다양한 데이터들을 같이 확인할 수 있으며, OWL과 RDF 등으로 구성되어 있는 LOD 데이터를 변환하여 검색 결과를 보여준다. 기존의 일반 데이터 검색 방식과 유사하다고 볼 수 있지만, LOD 데이터를 검색하기 위해서 특정 키워드를 입력하면 시맨틱 LOD 프레임워크와 동기화되어 동적 SPARQL 쿼리 생성 등의 과정을 통해 LOD 데이터 간의 관계(Relationship)를 기반으로 결과 데이터들과 연결된 일련의 데이터들을 같이 제공한다. 사용자는 이러한 데이터들의 연결성을 통해 더 확장된 정보를 확인하고 검색해나갈 수 있다는 점에서 차이점이 있다.

LOD 시각화 서비스는 그래프 형태와 문서 형태로 연관 데이터들을 시각화한다. LOD 시각화 서비스의 그래프 형태 표현 방식은 특히 [그림 4]의 (a)와 같이 서로 다른 도메인 간의 연결성을 표현하기 위해 도메인별 색상에 구분을 두어 정의하고 시각화한다. [그림 4]의 (b)는 그래프 표현을 위한 기본 요소들이 LOD 개체의 속성과 개체 간의 연결 관계 정보를 기반으로 설계된 정의서의 일부이다. LOD 시각화 서비스는 LOD 데이터 간의 연결 과정을 알고 싶은 사용자와의 상호작용을 지원한다. 그래프 형태의 각 LOD 개체를 선택할 수 있으며, 사용자가 개체를 선택하면 개체의 메타정보를 모두 보여준다. 또한, 검색한 LOD의 연결성을 알 수 있는 1레벨의 데이터뿐만 아니라 사용자가 발생시키는 이벤트 입력에 따라 N레벨까지의 연결 과정을 보여준다. 본 서비스는 실시간으로 업데이트되는 LOD를 즉각 반영하여 보여줌으로써 스마트시티 환경에서 발생하고 있는 주차 부족 구역 관리, 미세 먼지 수치를 반영한 새로운 서비스 창출 등으로 도시문제 해결을 가능하게 한다.

LOD 발행 서비스는 [그림 4]의 (c)와 같이 사용자

으로 제공하기 위해 동적으로 질의를 생성하고, 사용자와의 상호작용을 통해 데이터와 데이터 간의 의미 있는 연결성을 보여주며, 데이터 발행을 통해 도시 데이터의 효율적인 재사용성과 새로운 서비스 창출의 가능성을 강조한다. 서로 다른 도메인에서 수집되는 공공 데이터 간의 연결성을 검색 가능하게 함으로써 도시에서 실시간으로 일어나고 있는 도시 문제를 파악할 수 있으며 데이터를 활용하여 도시문제 해결을 위한 새로운 시스템을 도출할 수 있을 것이다. oneM2M에서는 이러한 시맨틱 기

능을 제공하고, 다양한 도메인에서 발행되는 데이터들을 통합하는 표준을 제공하기 위해, Rel-1에서 4까지 개발되었던 시맨틱 사물인터넷 기본 기능을 더욱 고도화하고 효율적인 검색이 가능하도록 신규 워크아이템을 Release 5의 주요 기능으로 선정하고 관련 표준 개발을 진행하고 있다. 향후 새로운 외부 온톨로지 도입과 LOD 데이터 학습 모델을 통해 스마트도시 내에서 발생하는 도시문제를 스스로 발견하고 해결할 수 있는 시스템에 대한 연구를 추진하고자 한다. 

※ 본 연구는 국토교통부/과학기술정보통신부/국토교통과학기술진흥원의 스마트시티 혁신성장동력 프로젝트 지원으로 수행되었음(과제번호 18NSPS-B149386-01).

참고문헌

- [1] Berners-Lee, T., Hendler, J., & Lassila, O. 'The semantic web.' Scientific american 284, no. 5 (2001), pp.34-43.
- [2] Berners-lee, T. 'Linked data.' (2006)
- [3] I. G. Luis, M. Ian, G. Hugh, and S. Elena, 'An assessment of adoption and quality of linked data in european open government data,' International Semantic Web Conference, pp. 1-17, June 2019.