

시맨틱 IoT 환경에서 그래프 순회를 위한 시맨틱 그래프 인터워킹 시스템

이지은^{1, 2}, 안종관^{1, 2}, 정승명², 송재승^{1*}

세종대학교 정보보호학과¹, 전자부품연구원 자율지능 IoT 연구센터²

{love9ly, cftn3212}@sju.ac.kr, sm.jeong@keti.re.kr, jssong@sejong.ac.kr*

Semantic Graph Interworking System for Graph Traversal in IoT Environment with Semantics

Lee JiEun^{1, 2}, An JongGwan^{1, 2}, Jeong SeungMyeong, Song JaeSeung^{1*}

Sejong University¹, Korea Electronics Technology Institute²

요 약

시맨틱 기술은 스마트시티의 사물인터넷 플랫폼을 통하여 수집된 수많은 원시 데이터에 의미를 부여하고 데이터 간의 관계를 형성함으로써 모든 데이터들이 유기적으로 연결될 수 있게 한다. SPARQL은 이러한 시맨틱 데이터를 검색할 수 있는 쿼리 언어로 사용되고 있지만, 검색 속도와 복잡한 연결관계 검색의 한계 등으로 인해 데이터 간의 순회가 가능한 그래프 데이터에 대한 요구가 높아지고 있다. 하지만 기존의 시맨틱 데이터와 그래프 데이터 간의 리소스 구조와 표현 범위의 차이로 인해 변환 과정에서 데이터가 손실되거나 불필요한 데이터가 생성되는 현상이 발생하게 된다. 따라서 본 논문에서는 스마트시티의 시맨틱 데이터를 손실되는 데이터 없이 그래프 순회가 가능하도록 표현하고 이를 기반으로 더 높은 수준의 지식을 도출할 수 있는 시맨틱 그래프 인터워킹 시스템을 제안한다.

I. 서 론

시맨틱 기술(Semantics)은 스마트시티에서 수집되는 데이터를 연결하여 데이터의 재사용 및 새로운 서비스 창출을 가능하게 한다. 시맨틱스는 Resource Description Framework (RDF), Web Ontology Language (OWL)과 같은 공통 어휘를 사용함으로써 시맨틱 데이터를 표현할 수 있으며, oneM2M의 Semantic Descriptor (SD) [1], FIWARE의 NGSI-LD(Linked Data) [2] 등 여러 국제 IoT 표준화 기구와 사물인터넷 오픈 플랫폼 그룹에서도 시맨틱 데이터의 리소스를 정의하여 스마트시티에서 사용될 수 있도록 표준화 활동이 이루어지고 있다. 하지만 시맨틱 데이터 전용 쿼리 언어인 SPAQL(Sparkle Protocol and RDF Query Language)은 데이터의 관계를 타고 들어가 연결된 특정 데이터 검색과 같은 복잡한 연결관계 검색에 어려움이 있으며 복잡한 구조일수록 검색속도가 현저하게 떨어지는 등의 한계가 있다. 반면 그래프 데이터는 복잡한 연결 관계도 빠르게 검색할 수 있는 그래프 순회가 가능하며 시맨틱 데이터와 흡사한 구조를 가지고 있어 스마트시티의 그래프 데이터에 대한 요구가 높아지고 있다.

하지만 IoT 표준별 시맨틱 리소스 구조의 차이와 기존의 시맨틱 데이터와 그래프 데이터 간의 데이터 표현 범위의 차이로 인해 변환 과정에서 데이터 손실 및 불필요한 데이터가 생성되게 된다. 따라서 본 논문에서는 IoT 표준의 시맨틱 데이터가 손실되는 데이터 없이 그래프 순회가 가능하도록 표현하고 이를 기반으로 더 높은 수준의 지식을 추출할 수 있는 시맨틱 그래프 인터워킹 시스템을 제안한다.

본 논문의 구성은 다음과 같다. 2 장에서는 시맨틱 데이터와 그래프 데이터 모델의 표현 범위를 비교함으로써 본 시스템의 필요성을 강조하고 시맨틱 그래프 인터워킹 시스템에 대해 설명한다. 마지막으로 3 장에서는 결론 및 향후 연구에 대해 논의한다.

II. 본론

본 단락에서는 시맨틱 데이터와 그래프 데이터 모델의 표현 범위의 차이에 대해 비교 분석하고, 시맨틱 그래프 인터워킹 시스템에 대해 서술한다.

2.1 데이터 모델 표현 범위

그림 1은 시맨틱 데이터와 그래프 데이터의 데이터 모델 개념과 표현범위를 비교하기 위한 그림이다. 본 논문에서는 시맨틱 데이터 중 하나로 FIWARE 그룹에서 제안하는 NGSI-LD의 데이터 모델 개념을 예시로 분석하였다. b)의 그래프 데이터는 Node (Vertex)가

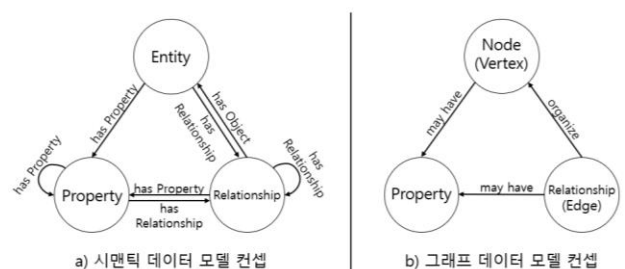


그림 1. 데이터 모델 표현 범위 : a) 시맨틱 데이터 모델 컨셉, b) 그래프 데이터 모델 컨셉

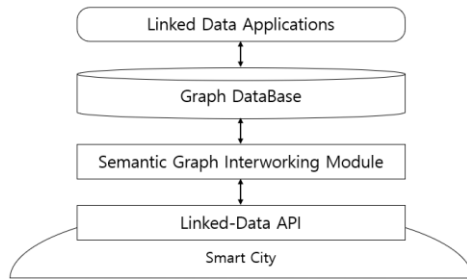


그림 2. 시맨틱 그래프 인터워킹 시스템 구조

Property 를 가지고 Relationship (Edge)으로 다른 Node 와 연결될 수 있는 시맨틱 데이터의 트리플 구조와 유사한 구조를 가지고 있다. 하지만 그래프 데이터와 다르게 a)의 시맨틱 데이터는 Property 가 Property 를, Relationship 이 Relationship 을 가지는 관계, Property 가 Relationship 을 가지는 관계를 추가적으로 표현할 수 있다. 이러한 표현은 일반적인 그래프 데이터의 표현 범위에서 벗어나는 형태이며, 그래프 데이터에 표현되지 못하는 리소스 구조는 시맨틱 데이터를 변환하여 그래프 데이터베이스에 저장할 때 손실되는 데이터와 불필요한 데이터를 생성하게 한다. 본 논문에서는 이러한 시맨틱 데이터의 데이터 손실 없이 시맨틱으로 기술된 데이터들을 그래프로 표현하여 연결함으로써 보다 다양한 그래프 검색 및 의미 추출을 가능하게 하는 시맨틱 그래프 인터워킹 시스템에 대해 제안하며 다음 단락에서 설명한다.

2.2 시맨틱 그래프 인터워킹 시스템

그림 2 는 본 논문에서 제안하는 시맨틱 그래프 인터워킹 시스템의 구조를 표현하는 그림이다. 스마트시티에 저장되어 있는 시맨틱 데이터는 oneM2M, NGSI-LD 등 스마트시티 내에서 사용되고 있는 표준 IoT 플랫폼 구조의 Linked-Data API 를 통해 시맨틱 그래프 인터워킹 모듈로 전달된다. 시맨틱 그래프 인터워킹 모듈로 인해 변환된 데이터들은 그래프 데이터베이스에 저장되고, 어떠한 복잡한 연결관계를 가진 데이터라도 그래프 순회를 통해 조회할 수 있게 된다. 또한 저장된 그래프로 표현된 시맨틱 데이터들은 높은 수준의 지식을 추출하여 스마트시티에서 일어나고 있는 각종 상황들을 감지할 수 있는 Linked Data Applications 에서 상황 분석의 기초로 사용될 수 있다.

인터워킹을 제공하는 시맨틱 그래프 인터워킹 모듈의 구성은 그림 3 에서 표현된다. Linked-Data API 로부터 전달받은 Linked-Data Source 는 사용된 IoT 표준별로 Linked-Data Retriever 에서 각각 포맷에 맞게 읽어온다. 시맨틱 데이터를 데이터 손실없이 변환하기 위해서는 변환할 그래프 데이터의 리소스 구조에 대한 기준을 잡기 위해 표준별로 시맨틱 데이터 리소스 구조에 대한 정보를 가지고 있는 Linked-Data Specification 가 필요하다.

시맨틱 그래프 인터워킹 모듈은 이러한 Linked-Data Specification 을 참조하여 그래프 데이터베이스에 저장될 수 없는 복잡한 구조를 가진 시맨틱 데이터를 변환하는 Graph Data Adaptor 를 포함하고 있다. Graph Data Adaptor 는 Linked-Data Specification 에 따라 그래프 데이터의 기준이 되는 Mapping Model 를 참조하여 시맨틱 데이터를 그래프 데이터로 변환한다. Graph Data Adaptor 가 Mapping Model 기반의 그래프 데이터를 전달해주면 Graph DB Writer 가 타겟 Graph

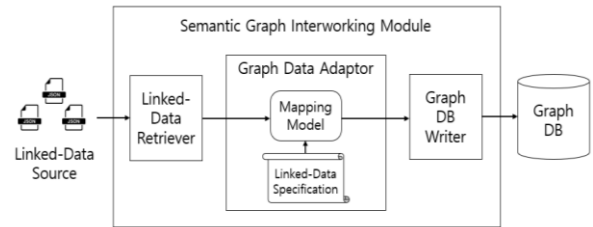


그림 3. 시맨틱 그래프 인터워킹 모듈 구성

DB 에 맞는 그래프 쿼리(e.g. AQL, Cypher)를 사용하여 그래프 데이터베이스에 적합한 데이터를 전달해줌으로써 인터워킹이 성공적으로 수행된다.

시맨틱 그래프 인터워킹 모듈은 데이터 변환과정에서 그래프 데이터베이스에 부합하지 않는 시맨틱 관계 정보들을 생략하지 않고 IoT 표준 시맨틱 데이터에 맞게 매핑 모델을 생성하여 변환함으로써 데이터 손실이 없는 온전한 시맨틱 데이터의 그래프 인터워킹을 가능하게 한다. 또한 그래프 순회가 가능하게 됨으로써 기존 SPARQL 의 한계를 극복할 수 있게 한다.

III. 시사점 및 결론

본 논문에서는 시맨틱 IoT 환경에서 그래프 순회를 위한 시맨틱 그래프 인터워킹 시스템을 제안하였다. SPAQL 쿼리 검색의 한계와 복잡한 연결관계의 그래프 순회를 위해 시맨틱 데이터의 그래프 데이터 변환이 요구되지만, 시맨틱 데이터와 그래프 데이터 사이의 표현범위의 차이와 리소스 구조의 차이 때문에 시맨틱 데이터의 일부 복잡한 관계가 생략되어 변환되는 경우가 발생한다. 본 논문에서는 기존 시맨틱 데이터를 데이터 손실 없이 그래프로 표현하여 시맨틱 데이터의 다양한 연결관계 검색 및 그래프 순회를 가능하게 하는 시맨틱 그래프 인터워킹 시스템을 제안하여 그래프 순회를 가능하게 한다. 이와 같이 그래프로 표현된 시맨틱 데이터들은 향후 보다 높은 수준의 지식 추출을 통해 스마트시티 환경에서 일어나는 다양한 상황 분석의 기반이 될 수 있을 것으로 기대한다.

ACKNOWLEDGMENT

본 연구는 국토교통부/과학기술정보통신부/국토교통과학

기술진흥원의 스마트시티 혁신 성장동력 프로젝트 지원으로

수행되었음(과제번호 18NSPS-B149386-01). 교신저자:

송재승 교수

참 고 문 헌

- [1] TS-0030 Ontology Based Interworking, release 3, v3.0.1, oneM2M.
- [2] Context Information Management (CIM); NGSI-LD API, ETSI GS CIM 009, v1.1.1, Jan. 2019.