

## A Formal Specification and Meta-Model for Development of Cooperative Collection·Analysis Framework

Eun-Sook Cho\*, Chee-Yang Song\*\*

\*Professor, Dept. of Software Engineering, Seoil University, Seoul, Korea

\*\*Professor, Dept. of Software, Kyungpook University, Daegu, Korea

### [Abstract]

Companies can identify user groups or consumption trends by collecting and analyzing opinions of many users on special subjects or their products as well as utilize them as various purposes such as predicting some specific trends or marketing strategies. Therefore current analyzing tools of social media have come into use as a means to measure the performances of social media marketing through network's statistical analysis. However these tools require expensive computing and network resources including burden of costs for building up and operating complex software platforms and much operating know-how. Hence, small companies or private business operators have difficulty in utilizing those social media data effectively. This paper proposes a framework applied into developing analysis system of social media. The framework could be set up and operate the system to extract necessary social media's data. Also to design the system, this study suggests a meta-model of proposed framework and to guarantee completeness and consistency, a formal specification of meta-model by using Z language is suggested. Finally, we could verify the clearness of framework's design by performing Z model checking of formal specification's output through Z-EVES tool.

▶ **Key words:** Social Media, Data Analysis, Framework, Meta-model, Formal Specification

### [요 약]

기업들은 소셜 미디어로부터 특정 주제 및 자사 제품에 대한 다수의 사용자들의 의견을 수집하고 분석하여 사용자 집단이나 소비 트렌드를 식별할 수 있고, 이를 기반으로 미래 예측 및 마케팅 등의 다양한 목적으로 활용하고 있다. 따라서 현재 소셜 미디어 분석 도구들이 네트워크 통계 분석을 통해 기업의 소셜 미디어 마케팅의 성과를 측정하기 위한 수단으로 사용되고 있다. 그러나 이러한 도구들은 방대한 양의 소셜 미디어 데이터를 수집하고 이를 분석하기 위해 고가의 컴퓨팅 자원 및 네트워크 자원을 소모하고, 복잡한 소프트웨어 플랫폼 구축 및 운용에 따른 비용 부담과 많은 운영 노하우를 필요로 한다. 그 결과 중소기업이나 개인 사업자의 경우는 이러한 소셜 미디어 데이터를 효율적으로 활용하지 못하는 어려움이 있다. 따라서 본 논문에서는 중소기업이나 개인 컴퓨터에서도 운영이 가능한 소셜 미디어 데이터 수집 및 분석 시스템 개발에 필요한 프레임워크 설계와 제시된 프레임워크 설계에 대한 완전성이나 일관성을 검증하기 위한 정형 명세 기법을 제시한다. 또한 정형명세 기법으로 Z 언어를 통해 명세한 결과를 Z-EVES Tool을 통해 Z 모델 체크를 수행하여 프레임워크 설계의 명확성을 검증한다.

▶ **주제어:** 소셜 미디어, 데이터 분석, 프레임워크, 메타모델, 정형 명세

- First Author: Eun-Sook Cho, Corresponding Author: Chee-Yang Song
- \*Eun-Sook Cho (escho@seoil.ac.kr), Dept. of Software Engineering, Seoil University
- \*\*Chee-Yang Song (cysong@knu.ac.kr), Dept. of Software, Kyungbuk University
- Received: 2019. 11. 15, Revised: 2019. 12. 16, Accepted: 2019. 12. 16.

## I. Introduction

한국인터넷진흥원에서 2018년에 발표한 '인터넷 이용 실태 조사' 자료에 따르면 그림 1과 같이 지난 10년 간 이용률은 15.0%p 증가('08년 76.5%→'18년 91.5%) 했으며, 연령별 이용률 격차가 대폭 감소하여 남녀노소 구분 없이 인터넷을 이용하고 있는 것으로 나타났다[1].

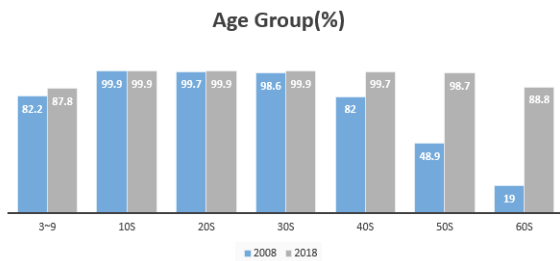


Fig. 1. Actual Condition Survey of Internet Utilizability

또한 소셜 미디어를 이용하는 사용자들은 다양한 목적에 의해 개인의 소소한 일상과 주관적인 의견, 전문적인 지식 등을 타인과 공유하는 커뮤니케이션의 수단으로 사용하고 있다. 그 결과, 기업들은 소셜 미디어(Social Media)로부터 특정 주제 및 자사 제품에 대한 다수의 사용자들의 의견을 수집하고 분석하여 사용자 집단이나 소비 트렌드를 식별할 수 있고, 이를 기반으로 미래 예측 및 마케팅 등의 다양한 목적으로 활용하고 있다[1,2].

현재 많이 사용되는 Expion이나 Gremin, Ubervu 등의 소셜 미디어 분석 도구들은 특정 단어에 대한 필터링과 특정 서비스를 통한 네트워크 통계 분석, 가시화 기능을 제공하고 있다[3,4]. 이런 도구들은 네트워크 통계 분석을 통해 기업의 소셜 미디어 마케팅의 성과를 측정하기 위한 수단으로 사용되며, 월 99 달러에서 1500달러까지의 사용 비용이 소요된다. 따라서 사회의 다양한 분야에서 이용현황 분석 및 향후 마케팅을 위해 방대한 양의 소셜 미디어 데이터를 수집하고 이를 분석하기 위해 고가의 컴퓨팅 자원 및 네트워크 자원을 소모하고, 복잡한 소프트웨어 플랫폼을 사용하여 데이터를 수집 및 분석하기 때문에 소프트웨어 플랫폼을 구축 및 운용하는데 비용이 많이 들뿐만 아니라 운영 노하우의 부족으로 소셜 미디어 데이터를 활용하는데 많은 어려움이 존재한다. 따라서, 본 논문에서는 이러한 문제를 해결하기 위하여 중소기업 및 개인 컴퓨터에 설치 및 운용이 가능한 스탠드얼론(Standalone) 형태의 소셜 미디어 데이터 분석도구에 적용할 수 있는 협력형 수집/분석 프레임워크 개발을 위한 메타모델과 이에 대한 검증으로 Z 언어[5]를 이용한 정형명세 기법을 제시한다.

본 논문의 구성으로 2장에서는 관련연구로 소셜 미디어 분석과 관련된 현황들에 대해 설명하며, 3장에서는 본 논문에서 제안하는 협력형 수집 및 분석 프레임워크 및 정형명세 기법을 제시한다. 4장에서는 3장에서 제안한 프레임워크 메타모델 정형 명세에 대한 검증결과를 제시하고, 마지막으로 5장에서 결론 및 향후 연구 과제를 제시한다.

## II. Preliminaries

### 1. Related works

#### 1.1 Analysis and Extraction of Social Big Data

SNS, 블로그, 뉴스 등 웹 페이지에서 크롤링이나 Open API를 이용해 수집된 정보들은 주로 비정형 데이터나 반(半)정형 데이터이다[7]. 데이터의 형태가 다양한 만큼 분석 기술 또한 전통적인 통계 알고리즘을 적용한 정형 데이터의 분석 기법과는 별개로 다양하고 복잡한 분석이 수행된다[8]. 비정형 데이터를 분석하기 위해서는 텍스트 분석을 시작으로 평판-감성 분석, 노출 추이 분석, 네트워크 분석 등이 기반이 되며, 특히 자연언어 이해를 위한 형태소 분석, 구문 분석, 개체명 인식, 이벤트 추출, 관계 분석 등의 자연언어처리 기술이 핵심이다[9,10]. 문서 키워드의 자동 분류와 군집, 요약 등이 수행되며 최근에는 자연어 처리를 포함해 기계학습과 딥 러닝(Deep Learning) 기술이 포괄적으로 활용되기도 한다[11]. 그러나 사용자 트렌드 분석과 마케팅을 위한 소셜 미디어 데이터의 수집 및 분석에 대한 많은 수요에도 불구하고, 다양한 소셜 미디어 서비스 인터페이스의 이질성으로 인한 동적 연동의 어려움과 소프트웨어 플랫폼 구축 및 운영의 복잡성을 해결하기 위한 연구는 미흡한 상태이다.

페이스북기의 분석 및 추출기법은 페이스북만의 특화된 소셜 토픽에 한정되어 추출하고 있는 기법이다. 또한 이 기법은 실시간 관심도가 높은 소셜 콘텐츠만 수집 대상이 된다. 이러한 시스템은 특정 소셜 채널에 특화된 소셜 토픽에 대해서는 적합하나, 여러 다양한 소셜 채널로부터의 데이터를 기반으로 범용적인 소셜 토픽 추출 및 분류를 목적으로 하는 경우에 있어서는 부적합하다는 한계점이 있다.

충북대[8]의 사용자 영향력 및 시간 변화를 고려한 소셜 네트워크 핫 토픽을 결정한 방법은 복수의 소셜 네트워크 콘텐츠에 포함되는 복수의 단어의 시간 슬롯의 변화에 따른 출현 빈도의 변화를 기반으로 단어를 추출하는 기법이다. 이 기법은 시간 슬롯 내에서 노출 빈도가 높은 단어에 대한 소셜 콘텐츠만이 대상이 되는 한계점이 있다. 본 논문에서는 보다 추출되는 소셜 토픽의 정확도를 높이기 위

해서 의미와 주제 연관성에 따른 분류를 통해 관련된 소셜 토픽을 추출하고자 한다.

본 논문에서는 이러한 문제를 해결하기 위해서는 동적 연동이 가능하면서 확장 가능한 소셜 미디어 콘텐츠 분석 시스템[11]을 구축하기 위한 프레임워크 개발이 필요하다. 프레임워크는 지속적으로 확장 가능하기 때문에 새로운 인터페이스들이 계속 추가되더라도 이를 반영할 수 있을 뿐만 아니라 여러 다양한 플랫폼 요구사항에 따라 커스터마이징 할 수 있기 때문이다.

### 1.2 Framework and Meta-Model

최근에는 대부분의 소프트웨어 개발 프로젝트에서 소프트웨어 개발의 생산성 향상과 재사용성 및 유지보수성을 향상시키기 위해 프레임워크 기반의 소프트웨어 개발 방식을 취하고 있다[12]. 프레임워크란 여러 유사한 애플리케이션들이 공통으로 필요로 하는 모듈로 구성된 뼈대로서, 하드웨어 메인 보드에 여러 재사용 부품인 메모리, CPU, 네트워크 카드, 그래픽 카드, 사운드 카드 등등의 컴포넌트들이 조합되어 있고, 이들 간에 서로 신호나 데이터를 주고 받을 수 있는 회로들이 설계 및 구현되어 있다. 이처럼 프레임워크 내에도 공통 모듈들이 재사용 가능한 컴포넌트 형태들로 구성되어 있으며, 컴포넌트들 간의 상호 작용을 위한 흐름이 설계되어 구현되어 있다. 본 논문에서는 시스템의 품질과 유지보수성 그리고 재사용성을 향상시키기 위해 소셜 미디어 콘텐츠 수집에서 분석, 추출 및 분류에 이르는 여러 다양한 컴포넌트들을 통합 및 관리하기 위한 프레임워크를 설계하고 제안된 프레임워크에 대한 정적 메타모델을 제시한다.

## III. The Proposed Scheme

이 장에서는 소셜 미디어 콘텐츠 분석 시스템의 핵심 엔진이라 할 수 있는 협력형 수집·분석용 프레임워크의 구조 및 각각의 구성 요소에 대해 설명한다[13]. 제시된 프레임워크는 소셜 미디어 콘텐츠 분석 시스템 개발에 적용 가능하며 기존의 복잡하고 고비용의 분석 시스템과 달리 인터페이스의 동적 연동 및 단순하면서 저비용으로 운영할 수 있도록 설계 한다.

### 1. Structure of Cooperative Collection·Analysis Framework

이 장에서는 소셜 미디어 콘텐츠 분석 시스템의 핵심 엔진이라 할 수 있는 협력형 수집·분석용 프레임워크의 구조

및 각각의 구성 요소에 대해 설명한다. 그림 2에 제시된 프레임워크는 크게 3가지 파트로 구분된다. 첫째는 소셜 미디어 콘텐츠를 수집하고 통합하는 부분이고, 둘째는 수집 및 통합된 소셜 미디어 콘텐츠들을 분석하고 분류하는 부분이고, 셋째는 수집되어 분석 및 분류된 소셜 미디어 콘텐츠들을 저장하는 부분이다. 협력형 수집·분석 연계 프레임워크를 구성하는 구성요소들로는 ‘소셜 컨텍스트 수집기’와 ‘의미 연관성 기반 소셜 토픽 추출기’, ‘주제 연관성 기반 소셜 콘텐츠/피드백 분류기’, ‘딥러닝 수행기’, ‘협력형 통합 지식베이스 관리기’로 구성된다.

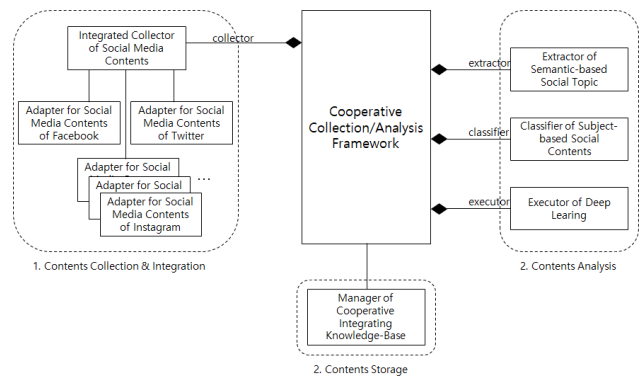


Fig. 2. Structure of Framework

제시된 프레임워크는 소셜 미디어 콘텐츠 수집, 분석, 추출, 분류의 과정을 지속적으로 운영하기 위해서 필요(Required) 또는 제공(Provided)하는 인터페이스들을 그림 5와 같이 설계하고 있다. 소셜 컨텍스트 수집기에는 여러 다양한 소셜 미디어 채널들로부터 소셜 미디어 콘텐츠들을 수집하기 위해서 소셜 미디어 모니터에서 제공하는 통합 API를 개발하여, 트위터(Twitter), 페이스북(Facebook), 인스타그램(Instagram) 등과 같이 계속해서 늘어나는 다양한 소셜 미디어의 오픈API를 활용하여 연동할 수 있는 수집 인터페이스(‘CollectIF’)를 그림 3과 같이 제공한다. 프레임워크에서는 수집 인터페이스를 통해 수집된 소셜 미디어 콘텐츠들을 의미 연관성과 주제 연관성을 토대로 분석하여 시스템이 학습할 수 있는 온톨로지(Ontology) 형태로 데이터 모델을 구축하여 스토리지에 저장한다. 이에 대한 인터페이스로 분석 인터페이스(‘AnalysisIF’)와 서비스 인터페이스(‘ServiceIF’)를 제공한다. 또한 수집기로부터 수집되어 분석되어 스토리지에 저장된 소셜 미디어 컨텍스트(Context)들을 관심 주제 혹은 관심 시점에 따라 동적으로 분류하고 관리하기 위해서는 지속적으로 딥 러닝을 통해 미디어 콘텍스트의 주제 추출을 해야 한다. 이를 위해서 실행 인터페이스(‘Executef’)가 요구된다.

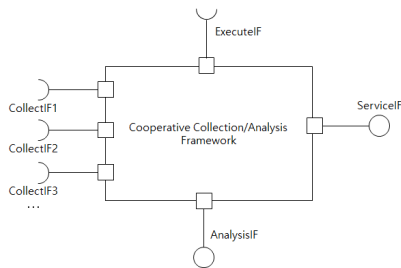


Fig. 3. Interfaces of Framework

제시된 프레임워크의 인터페이스들과 내부 구성 요소들 간의 관계 및 구조는 그림 4와 같다.

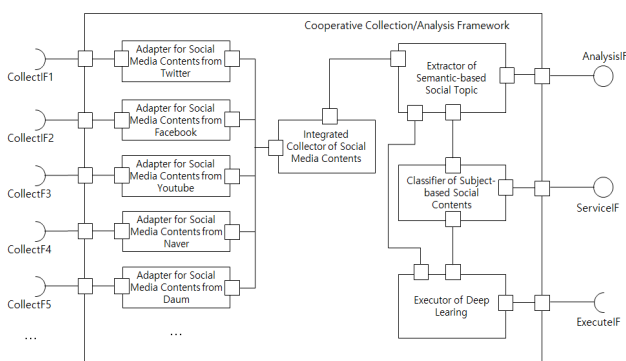


Fig. 4. Internal Structure of Framework

협력형 수집 및 분석 프레임워크의 내부 구성 요소들에 대한 특징은 다음과 같다.

• 소셜 미디어 콘텐츠용 어댑터

본 논문에서는 소셜 미디어 콘텐츠 채널의 소셜 컨텍스트 수집 API 이질성 문제와 추가적 유지보수 비용 문제를 효과적으로 해결하기 위해서 그림 5와 같이 어댑터(Adapter) 패턴을 적용한다.

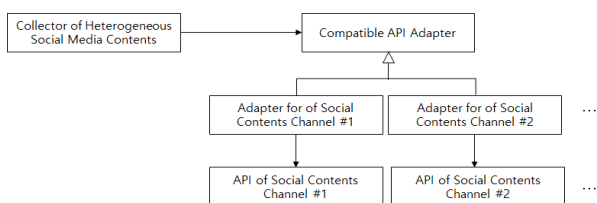


Fig. 5. Adapter Pattern for Collector of Social Media Contents

이 기법을 적용하면, 다양한 소셜 콘텐츠 채널의 API에 의존하지 않고, API 호환 어댑터를 통해 소셜 컨텍스트를 수집할 수 있다는 장점이 있다.

• 소셜 미디어 콘텐츠 통합 수집기

소셜 미디어 서비스는 사용된 인터페이스 언어와 제작

사에서 정의한 독자적인 인터페이스 구조에 따라 이질적인 인터페이스를 제공하기 때문에 단일 소프트웨어를 통해 모든 소셜 미디어 서비스로부터 데이터를 수집하는 것은 어렵다. 즉, 페이스북이나 트위터 등 각 소셜 미디어 서비스에서 제공되는 인터페이스 규격에 맞춘 수집 도구가 필요하다.

• 의미 기반 소셜 토픽 추출기

소셜 토픽은 시간 슬롯의 변화에 따라 노출 빈도가 높은 단어를 기반으로 추출된다. 소셜 미디어 콘텐츠는 비정형 데이터로 다수의 단어를 포함하고 있고, 단어의 노출 빈도만으로 소셜 토픽을 추출하는 것은 제한적이다. 따라서, 소셜 미디어 콘텐츠의 의미 정보를 포함한 소셜 컨텍스트를 기반으로 토픽 추출을 진행한다. 소셜 컨텍스트의 의미 연관성은 콘텐츠를 구성하는 단어들 간의 연관 관계를 기반으로 학습이 될 때, 보다 높은 정확도의 소셜 토픽 추출이 가능하다.

• 주제 기반 소셜 콘텐츠 분류기

소셜 토픽 추출이 완료되면, 각 토픽에 관련된 소셜 콘텐츠 및 그에 대한 피드백을 분류하여 사용자의 접근성이 높게 유지해야 이용편의성이 높아진다. 소셜 토픽은 소셜 컨텍스트의 의미 연관성을 기반으로 추출되기 때문에 이를 이용하여 각 소셜 콘텐츠와 피드백의 주제를 식별하여 분류한다. 제시하는 프레임워크에서는 의미 연관성 기반 소셜 토픽을 이용한 분석 대상 데이터(콘텐츠/피드백) 분류 기술과 마이닝 기법 기반 콘텐츠 주제 추출 및 딥 러닝 기반 인덱싱 기술을 이용하여 분류하도록 한다.

• 딥 러닝 학습기

딥 러닝 학습기는 소셜 미디어 컨텍스트를 관심 주제 혹은 관심 시점에 따라 동적으로 분류 및 관리하기 위해 필요하다. 사용자의 관심 주제와 관심 시점은 고정되지 않고, 사용자에 따라 서로 다르기 때문에 효율적인 미디어 컨텍스트 관리를 위해서는 딥 러닝을 통해 소셜 미디어 컨텍스트의 주제 추출을 지원한다. 딥 러닝 학습기는 소셜 토픽 추출 및 소셜 콘텐츠·피드백 분류를 하기 위해 단어 간 의미 연관성 분석과 주제 연관성 분석을 위한 딥 러닝 학습 데이터를 수집하고 정제한다. 딥 러닝 실험의 초기에는 지도 학습(Supervised Learning)으로 진행되며, 다양한 소셜 미디어 채널로부터 일정한 기준에 따른 학습 데이터를 수집한다.

## 2. Meta-Model of Cooperative Collection·Analysis Framework

그림 4의 협력형 수집 및 분석 프레임워크 설계를 기반으로 하여 이 프레임워크에 대한 메타모델을 그림6과 같이 정의하였다. 그림 6에 제시된 바와 같이 재사용 프레임워크는 통합 수집기(Integrated Collector), 의미 기반의 소셜 토픽 추출기(Extractor of Semantic-based Social Topic), 주제 기반의 분류기(Classifier of Subject-based Social Contents), 딥러닝 실행기(Executor of Deep Learning), 그리고 협력형 통합 지식 베이스 관리기(Manager of Cooperative Integrating Knowledge Base)로 구성된다. 각각의 소셜 미디어 채널 별 어댑터는 통합수집기가 포함하고 있다.

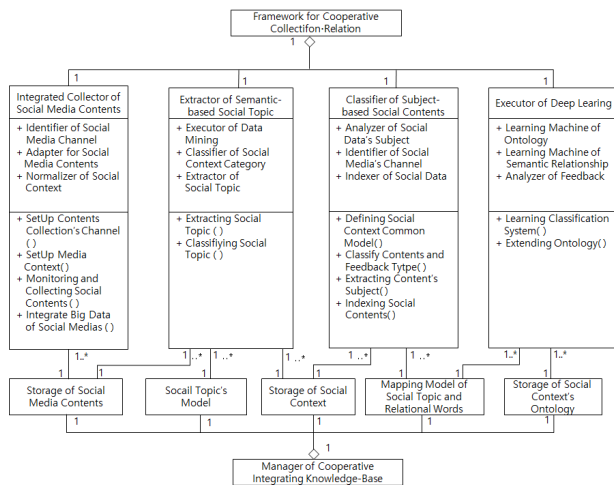


Fig. 6. Meta-model of Framework

그런데 제시된 프레임워크를 Z로 명세하고 검증하기 위해서는 그림 6의 클래스명을 간명하게 부여하고, 속성과 연산을 제거하고, 클래스간의 관계를 새로이 부여하여 작성한 것이 그림 7이다.

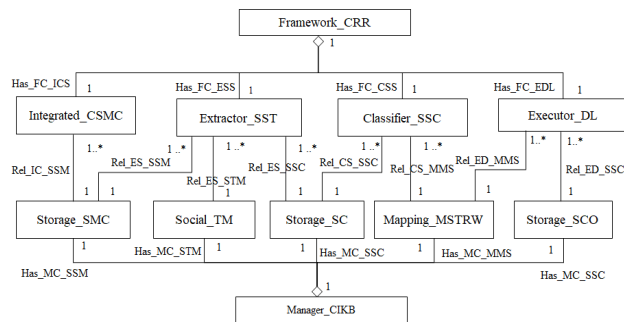


Fig. 7. Simple Meta-model of Framework

그림 7에서 협력형 수집 및 분석 프레임워크(Framework\_CRR)는 Integrated\_CSMC, Extractor\_SST, Classifier\_SSC, Executor\_DL로 구성하며, 협업 통합 관리

기(Manager\_CIKB)는 Storage\_SMC, Social\_TM, Storage\_SC, Mapping\_MSTRW, Storage\_SCO의 요소를 포함하는 관계로 되어 있다.

## 3. Formal Specification of Cooperative Collection·Analysis Framework

조은숙[13]과 송치영[14]에 제시된 클래스 모델의 Z로의 변환 규칙을 이용해서 그림 8의 프레임워크 메타모델을 Z로 정형화하여 명세한다. 메타모델 내에 있는 클래스 요소에 대해 기본 타입과 free type들을 선언하고, 또한 이들에 대한 스키마를 정의하고, 그리고 관계상에 존재하는 각 relationship 별로 스키마 정의를 한다. 이때, static semantics들을 위한 invariant한 constraint들은 요소를 정의한 요소와 관계 기반을 통해 생성된 스키마의 술어부에 명세한다.

Z 명세의 구축 결과로서, 기본타입으로 Storage\_SMC 외 10개의 기본 타입을 선언하고, 요소 기반 스키마 명세로 Framework\_CRR의 10개를 명세하고, 그리고 관계 기반 스키마 명세로 Has\_FC\_ICS의 16개를 작성하였다.

본 절에서는 구축된 Z 명세를 기술함에 있어 일부를 보인다. 이러한 Z의 정형적 명세를 통해서, 프레임워크 메타모델의 의미를 명확하게 할 수 있으며, 이를 통해 구축 모델에 대한 속성들(구문적 일관성 등의 Z/EVES 통한 메타모델의 구문적 분석 및 검증이 가능해진다.

### • 기본타입 선언

기본 타입의 선언은 프레임워크 메타모델에 포함된 모든 요소들에 대해서 각 요소별로 기본 타입을 명세한다. 기본 타입의 선언은 다음과 같다.

[Storage\_SMC\_ID, Social\_TM\_ID, Storage\_SC\_ID, Mapping\_MSTRW\_ID, Storage\_SCO\_ID, Manager\_CIKB\_ID, Integrated\_CSMC\_ID, Extractor\_SST\_ID, Classifier\_SSC\_ID, Executor\_DL\_ID, Framework\_CRR\_ID]

### • 스키마 명세

Z의 스키마 명세는 협업 프레임워크 메타모델에 포함된 모든 요소와 관계들에 대해서 Z 스키마로 명세한다. 즉, 요소 기반의 스키마와 관계 기반의 스키마로 Z 명세를 작성한다. 요소 기반의 Z 스키마 명세는 메타모델 내 한 개 요소(클래스)에 대해 한 개의 Z 스키마로 정의한다. 이때, 요소들 간에 복잡한 구조를 갖는 스키마 정의의 순서는 상향식 접근(Bottom-up Approach)으로 하부 요소에서 상위 요소를 대상으로 Z 스키마를 명세해 나간다[2]. 그래서, association, aggregation,

and composition의 관계를 갖는 요소들에 대해, 하부 구성 요소들을 먼저 명세하고, 점차 상위 요소들로 명세해 나간다. association 관계의 경우는 두 요소에 대해 순서에 상관없이 각각 스키마로 정의해주면 된다. 반면, aggregation과 composition 관계의 표현은 두 요소에 대해 각각 스키마로 명세한 후, 전체(whole)요소의 스키마 내 시그네처 부에 부분(part) 요소의 스키마를 포함하여 명세한다. 예를 들어, 그림 8에서 'Manager\_CIKB' 클래스는 'Storage\_SMC' 등 5개와 Aggregation의 관계를 갖는다. 이에 대한, Z 스키마 명세는 'Storage\_SMC' 등의 부분 객체들을 먼저 스키마로 정의하고 'Manager\_CIKB' 전체 객체의 스키마 정의시 부분 객체들을 내부 요소로 시그네처 부에 기술한다.

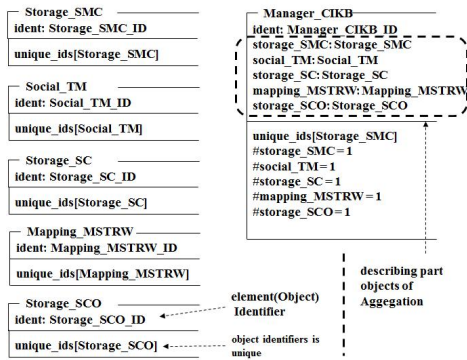


Fig. 8. Formal Specification based on Elements of Cooperative Framework's Meta-model

관계 기반의 Z 스키마 명세는 두 클래스 요소간의 관계 별로 각각 스키마로 정의한다. 이 스키마 명세는 클래스 간 관계 유형에 맞게 그리고 다중성(multiplicity)을 명확하게 구분하여 정의할 수 있다. 협력형 수집 및 분석 프레임워크 메타모델의 관계 기반 스키마 명세 17개 중에서 일부를 나타낸 것이 그림 9이다.

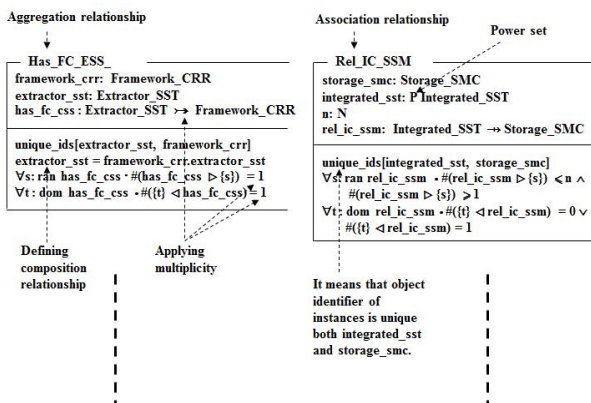


Fig. 9. Formal Specification based on Relations of Cooperative Framework's Meta-model

그림 9에서, 협력형 수집 및 분석 프레임워크 메타모델에서 관계의 명칭 부여는 연관(association)은 Rel\_로, 집합(aggregation)은 Has\_로 약칭하였다. 그림 10에서 각 관계의 유형별, 즉, 연관 및 집합에 대해 그 명세를 보여주고 있다. 관계 스키마의 시그네처 부에는 각 relation별로 연결된 요소들과 이들 간의 함수관계를 선언한다. 술어부에는 다중성(multiplicity) 및 불변 제약사항을 기술한다. 시그네처 부에서는 다중성에 따라 두 요소 간 매핑 함수가 상이하게 정의하며, 술어부에는 클래스간의 다중성을 각각 명확하게 기술한다.

먼저, 집합 관계에 대해서, 가령 'Has\_FC\_ESS' 집합 관계에서, 이 관계에 관련되는 요소는 Framework\_CRR와 Extractor\_SST이다. 시그네처 부에는 관계하는 두 요소인 Framework\_CRR와 Extractor\_SST를 선언하고, 다중성 1:1에 기반 하여 두 요소가 갖는 함수관계 bijection으로 명세한다. 술어부에서, 첫 줄 unique\_ids 명세는 "Extractor\_SST"와 "Framework\_CRR"의 instance인 extractor\_sst와 framework\_crr가 유일하다는 것을 표현한다. 두번째 줄은 집합(aggregation) 관계에 대해서 "Extractor\_sst = framework\_crr.extractor\_sst"로 명세한다. 이하 줄은 domain과 range에 restriction을 주어 두 요소간의 cardinality인 1:1을 표현한다.

다음으로, 연관 관계를 갖는 'Rel\_IC\_SSM' 관계에서 관계하는 요소는 Storage\_SMC와 Integrated\_SST이다. 연관 관계는 스키마의 시그네처 부에 두 요소를 각각 변수로서 선언하면 된다. 이 두 요소는 요소 기반 스키마 명세에서 이전에 스키마로 정의되었다. 이 요소에 대한 변수 선언 시, 다중성에 의해 단일 set(객체가 1개) 혹은 Power set(객체가 2개 이상)으로 표현된다.

#### IV. Experiments and Evaluation

본 논문에서 제시한 프레임워크와 이에 대한 메타모델의 구문적 및 의미적인 적합성을 검증하기 위해서 Z/EVES Tool을 사용하여 Z 명세의 구문검사, 변수 범위 검사 및 일관성 검사를 위해 Z/EVES Tool을 통하여 모델을 체크링(checking)한다[13]. 그 이유는 프레임워크 메타모델은 Z 스키마로 변환 명세 되었고, 명세의 구문적 및 정적 시맨틱 속성이 정확하다고 검사된다면 결국, 프레임워크 메타모델이 명확하다는 것을 뜻하기 때문이다. 이 툴에 Z 명세한 것을 입력하고 검사를 실행하여 syntax 및 type checking, 그리고 domain checking을 수행한다. Z/EVES 툴을 통한 검사

결과로서, 그림 10은 11개 타입 선언에 대한 명세에 대한 검사 결과를 보여준다. 그림 10에서 상단의 syntax와 proof 메뉴를 보면, 그 이하 부문에 각각 "Y"로 실행 결과가 나온 것은 Z의 기본 타입의 선언에 syntax와 proof 속성(property)에 대해 오류가 없음을 확인할 수 있다.

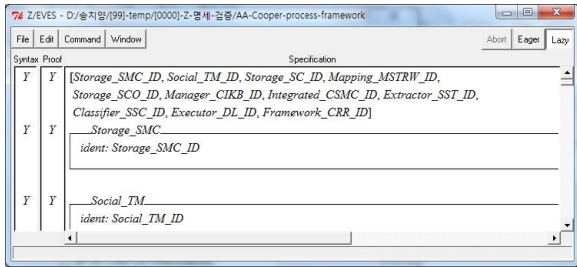


Fig. 10. Result of Syntax Verification for Basic Type's Z specification

다음으로 그림 11은 그림 12의 요소 기반 Z 스키마 명세인 11개에 대한 모델 검사의 결과를 보여주고 있다. 그림 11에서 왼쪽 상단에 표시된 'syntax'는 syntax와 타입 체크를 보여주는 것으로 'Y'로 나타나면 타입 간 일치성이 있음을 나타낸다. 반면, 'proof'는 도메인 체크의 결과를 보여주는 것으로 'Y'로 나타나면, 그 명세가 정확함을 의미한다.

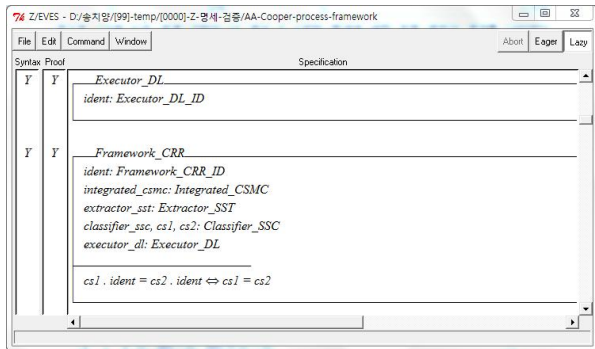


Fig. 11. Result of Syntax Verification for Element-based Z specification

끝으로, 그림 12는 관계 기반 Z 스키마 명세인 17개(집합 관계 9개, 연관관계 8개)에 대한 모델 체크의 결과를 나타낸다. syntax와 proof 부문이 "Y"의 결과를 보임으로서 명세가 정확함을 알 수 있다. Proof는 reduce에 의해 prove하였다(prove by reduce). 결국, 이와 같은 모델 체크를 통해, 협력 프레임워크 메타모델의 구문적 및 정적 시맨틱 측면에서 오류가 없음을 입증할 수 있다. 즉, 클래스간의 관계와 대응 수를 명확하게 명세하여 검사하여 메타모델의 구문적 구조에서 요소들 간 불일치 등의 error 혹은 conflict가 발생하지 않았다.

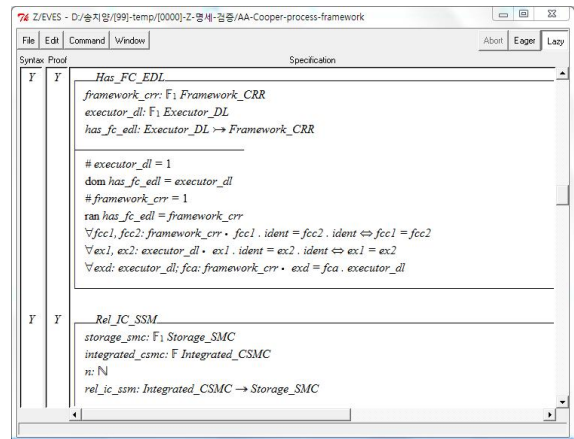


Fig. 12. Result of Syntax Verification for Relation-based Z specification

## V. Conclusions

중소기업 및 개인 사업자 등은 소셜 미디어 데이터의 수집 및 분석에 대한 필요성을 인식하고 있지만, 이를 위해 복잡한 소프트웨어 플랫폼을 구축 및 운용하기 위해 필요한 비용 부담과 운영 노하우의 부족으로 소셜 미디어 데이터를 활용하지 못하는 문제점이 존재한다. 본 논문에서는 이러한 문제점을 해결하기 위해서 재사용 가능하면서 동적으로 간단하게 연동할 수 있는 협력형 소셜 미디어 콘텐츠 수집 및 분석을 위한 프레임워크에 대한 설계 기법과 제시된 설계기법을 제시했다.

아울러, 제시하는 협력형 수집 및 분석 프레임워크의 메타모델에 대한 Z 언어를 이용한 정형적 명세와 Z/EVES를 사용한 모델 체크를 통해 이 메타모델이 정확함을 입증하였다. 이를 통해 중소기업이나 개인 사업자도 소셜 미디어 분석과 소셜 마케팅 통합 분석 서비스를 효율적으로 수행할 수 있다. 또한 현재 제시된 설계를 기반으로 여러 다양한 소셜 미디어 분석 및 추출 시스템 개발에 활용은 하지 못했다.

향후 연구에서는 제시된 프레임워크를 적용한 여러 다양한 소셜 토픽 추출 시스템을 개발하기 위한 개발 프로세스와 적용 사례를 연구할 것이다.

## ACKNOWLEDGEMENT

The present research has been conducted by the Research Grant of Seoil University in 2019.

## REFERENCES

- [1] KISA(Korea Internet & Security Agency), 2018 Survey on the Internet Usage, <https://www.msit.go.kr>
- [2] Byoung-Yup Lee, Jong-Tae Lim, Jaesoo Yoo, "Utilization of Social Media Analysis using Big Data", Journal of the Korea Contents Association, Vol.13, No. 2, pp.211-219, Feb., 2013.
- [3] Man-Mo Kang, Sang-Rak Kim, Sang-Moo Park, "Analysis and Utilization of Big Data", Communications of the Korean Institute of Information Scientists and Engineers, Vol. 30, No. 6, 2012.6, pp. 25-32, June, 2012.
- [4] Keun-Tae Kim, "Environment Challenge in Company for Big Data Analysis", Korea Information Processing Society Review, Vol.19, No.2, March, 2012.
- [5] M. Saalink, *The Z/EVES 2.0 Users Guide*, TR-99-5493-06A, ORA Canada, 1999.
- [6] Kee-young Kwahk, "Social Network analysis", The Second Edition, Chungnam, 690 pages, 2017.
- [7] Sang-hyeok Park, Seung-hee Oh, Haeng-nam Sung, "Analysis of the Usage patterns of Social Network Service Users", Journal of the Korea Society of Digital Industry and Information Management, Vol.9, No.4, pp. 251-265, 2013.
- [8] Yeon-woo Nho et al., "Reliable Hot Topic Prediction Scheme Considering User Influences in Social Networks", Journal of the Korea Contents Association, Vol.15 No.8, Aug. 2015.
- [9] Kie-jin Park, "A Design on Informal Big Data Topic Extraction System Based on Spark Framework", KIPS Transaction of Software and Data Engineering, Vol.5, No.11, pp.521~526, 2016.
- [10] Jin-myeong Chung, Young-ho Park, Woo-ju Kim, "Social Media Analysis Based on Keyword Related to Educational Policy Using Topic Modeling", Journal of Korean Society for Internet Information, No.19, Vol. 4, pp.53-63, Aug. 2017.
- [11] Eun-sook Cho, et. al., "Development of Extracting System for Meaning-Subject Related Social Topic using Deep Learning", Journal of the Korea Society of Digital Industry and Information Management 2018, vol.14,no.4,pp. 35-45 (11 pages)
- [12] Njeru Mwendu Edwin, "Software Frameworks, Architectural and Design Patterns", Journal of Software Engineering and Applications, JSEA Vol.7 No.8, pp.670-678, July 2014.
- [13] Eun-sook Cho, Chul-jin Kim, Chee-yang Song, "Formal Specification of Reusable Framework of Embedded System", Journal of KIPS D, Vol. 17-D, No. 6, pp.431-442, December, 2010.12.
- [14] Chee-yang Song, "A Meta-model-Based Modeling Mechanism for Hierarchical Design in UML", Thesis for the Degree of Doctor, 2003. 7.

## Authors



Eun-Sook Cho received the B.S. degree in Computer Science from DongEui University, Korea in 1993. He received the M.S and Ph.D degree in Computer Science from SoongSil University, Korea, in 1996 and 2000, respectively. Dr. Cho joined the faculty of the Department of Software Engineering at Seoul University, Seoul, Korea, in 2005. He is currently a Professor in the Department of Software Engineering, Seoul University. He is interested in component-based Development, cloud computing, and IoT Applications.



Chee-Yang Song received the B.S. degree in Computer Science from Hannam University, Korea in 1985. He received the M.S degree in Computer Science from Chung-Ang University, Korea, in 1987, and received the Ph.D degree in Dept of Computer Science from Korea University, Korea, in 2003, respectively. Dr. Song joined the faculty of the Department of software at Kyungbuk University, Sangju, Korea, in 2005. He is currently a Professor in the Department of Software, Kyungbuk University. He is interested in component-based Development, cloud computing, and IoT Applications.