



2022 한국소프트웨어종합학술대회 Korea Software Congress 2022

2022. 12. 20.(화)~23.(금)
라마다 프라자 제주호텔

“반도체 시대를 선도할 정보기술”



12.20

[특별세션]

- 우수논문발표세션

[워크샵]

- 디지털 플랫폼 도시, 성과와 도전과제 워크샵
- 대용량 데이터 분석 최신 기술 및 데이터베이스 교육 워크샵
- 서울 협업 행동 지능 핵심기술 워크샵
- 전원 불안정 감내 자율 에너지 기반 컴퓨팅 워크샵
- GEdge Platform 커뮤니티 제5회 컨퍼런스
- 개방형 확장현실 플랫폼 기술개발 워크샵

12.22

[기초강연]

- 이재진 교수(서울대학교)
- 장우석 Master(삼성전자)

[초청강연]

- 김상연 교수(한국기술교육대학교)
- 전원석 책임연구원(현대자동차)
- Ph.D. Yuichi Nakamura(NEC Corp.)
- 이상원 교수(성균관대학교)
- 김성진 이사(카카오)
- 노강호 상무(삼성전자)

12.21

[특별세션]

- Top Conference 세션
- 우수논문발표세션

[워크샵]

- 한국소프트웨어공학역사워크샵
- Kakao Tech Workshop
- 연구지원기관 사업설명회
- Emerging Technology Workshop in Computer Systems
- 부산대학교 인공지능융합연구센터 성과공유 워크샵
- 2022 AI 데이터 품질 개선 오픈랩 프로그램
- 2022년 예지형 시각지능 과제 워크샵
- MEC, 연합학습, 블록체인 융합 기술 워크샵
- 인하대학교 AI 확산 및 산학협력 전략 워크샵
- 엣지마이크로데이터센터 시스템SW 워크샵
- 차세대융합기술연구원 경기도자율주행센터 성과 발표회

[특별세션]

- 박사포럼
- Top Conference 세션
- 우수논문발표세션

[워크샵]

- 코드 분석과 오류 마이닝이 결합된 SW 오류 자동 수정 기술 개발
- 융합 서비스를 위한 인공지능 & 모바일 기술 활용 워크샵
- 강원권 SW중심대학 협력워크샵
- 빅데이터 엣지클라우드 서비스 워크샵

12.23

[특별세션]

- Top Conference 세션
- 우수논문발표세션

[워크샵]

- 초거대 그래프의 지능적 고속 처리를 위한 그래프 DBMS 기술 개발 워크샵

동사 유사성 식별 기반 요구사항부터 유스 케이스 생성

저자1 윤예동^o 저자2 장우성 저자3 문소영 저자4 김영철

저자 소속 홍익대학교

저자1 yedong@mail.hongik.ac.kr, 저자2 uriel200@hongik.ac.kr, 저자3 whit2@hongik.ac.kr, 저자4

bob@hongik.ac.kr

Use case generation from Requirement Specifications based on Verb Similarity Mechanism

저자1 Yedong Yoon^o 저자2 Woo Sung Jang 저자3 So Young Moon 저자4 Robert Youngchul Kim

저자 소속 Hongik University

요 약

프로젝트의 성공을 위해, 요구사항을 올바르게 파악하는 것은 매우 중요하다. 특히 비정형화 요구사항의 분석을 통한 상위 수준 모델의 추출은 매우 중요한 이슈이다. 요구사항으로부터 상위 모델 추출을 위해, 본 논문은 자연어 기반 요구사항의 유스케이스 추출 메커니즘을 제안한다. 1) 각 자연어 요구사항을 구문 분석하여 기본 단위 품사를 식별한다. 2) Fillmore의 프레임 의미(Frame Semantic) 분석을 통해 주격 동사, 명사들을 가시화 모델화(Visual Modeling)한다. 3) 요구사항 내 동사들의 유사성 분석을 위한 Loosely connected 그래프 메커니즘을 제안한다. 결과적으로 자연어 요구사항에서 유스케이스를 추출할 수 있다. 이는 객체 지향 소프트웨어 개발에서 Use case Point(UCP) 기반 비용 예측을 위한 기반이 될 수 있다.

1. 서 론

요구사항 단계에서 발생한 오류를 개발 단계 또는 테스트 단계에서 수정하려면 많은 비용이 발생하기 때문에 요구사항의 올바른 분석은 매우 중요하다. 그럼에도 불구하고 요구사항의 잘못된 분석, SW 개발 시간의 부족은 현재까지도 소프트웨어 개발 과정에서의 문제 중 가장 큰 문제로 지목되고 있다[1,2].

이를 해결하기 위해 많은 정형 명세 기법이 제시되었다. 대한민국은 2012년 소프트웨어 산업 법제화 이후 요구사항 명확화를 명시하였다. 하지만 이러한 노력에도 불구하고 실제로 작성된 자연어 비정형 요구사항의 수준은 만족스럽지 못한 수준이다[3].

본 논문은 기존 문제의 해결을 위해 상위 문서에 정의된 문장 내에서 중요한 단어가 나타내는 문맥적 의미를 가시화하고, 유스케이스를 추출하는 방법 제안한다. 이 방법은 비정형 자연어 요구사항 문장들의 유사성을 식별할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서 Fillmore의 Frame Semantic과 비정형 요구사항으로부터 유스 케이스를 도출 방법을 언급한다. 3장에서 자연어 비정형 요구사항으로부터 Frame Semantic을 통해 유사어 식별하고 이를 가시화 모델링을 통해 Use Case를 추출하는 방법을 언급한다. 4장에서 결론을 언급한다.

2. 관련 연구

2.1 Fillmore의 Frame Semantic

Fillmore의 Frame Semantic[4]은 문장 내 구성요소들 간의 의미적(semantic), 구문론적(syntactic) 조합이다.

Frame은 다양한 참가자, 속성 및 기타 개념적 역할을 포함하는 상황을 의미한다. Frame은 Lexical Unit(LU)과 Frame elements(FEs)를 원소로 포함한다. LU은 Frame에 포함될 수 있는 단어들의 집합이다. FEs는 서로 다른 Frame들 간의 관계에 대한 집합이다.

2.2 비정형 요구사항으로부터 UCP 기반 비용 예측 방법

비정형 요구사항으로부터 UCP 기반 비용 예측 방법[5]은 Refined Fillmore's case grammar를 기반으로 요구사항 문장으로부터 Use Case Point(UCP)를 구하여 요구사항의 비용을 예측한다. 하지만 이 방법은 단어의 유사성을 고려하지 않는다.

그림 1은 Fillmore's case grammar 를 객체지향 개념에 맞게 재정의한 Refined Fillmore's case grammar이다. 이것은 문장 구성요소들의 격(case)을 Actor, Counter actor, Object, Theme, Result, Source, Instrument, Experience, Goal로 구분한다.

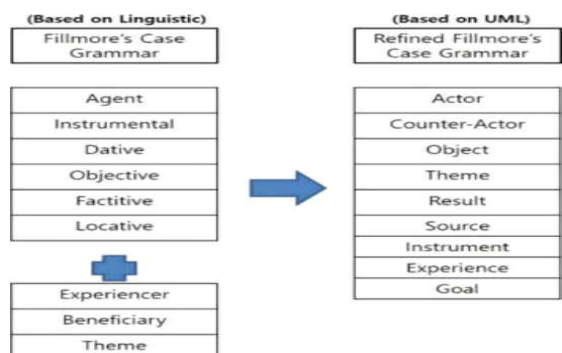


그림1 Refined Fillmore's case grammar

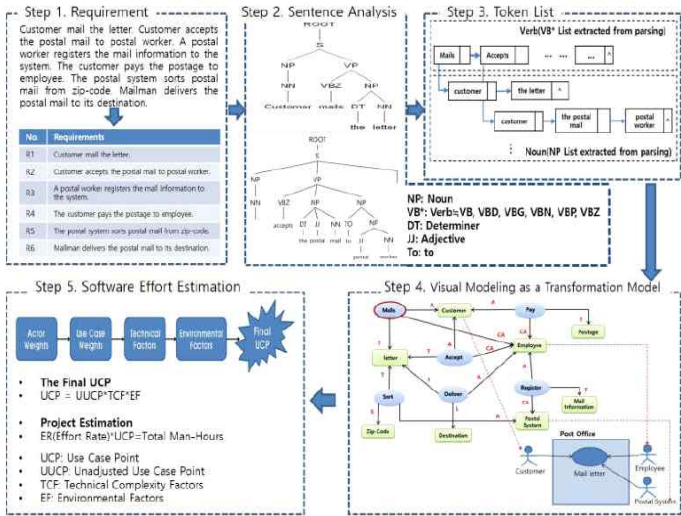


그림2 비정형 요구사항으로부터 UCP 도출 절차

그림 2는 비정형 요구사항에서 유스케이스 도출 후 UCP 계산 절차이다[5]. 자세한 절차는 다음과 같다.

1. 목표를 중심으로 비정형 자연어 요구사항을 그룹화한다.
2. 자연어 요구사항 문장을 분석하여 품사를 태깅한다.
3. 태깅된 품사에서 동사와 명사를 리스트화 한다.
4. Refined Fillmore' s case grammar를 기반으로 리스트 내 요소들의 격을 식별한다. 리스트를 기반으로 가시화 모델을 생성하고, 유스케이스를 추출한다.
5. 유스케이스로부터 UCP를 측정한다.

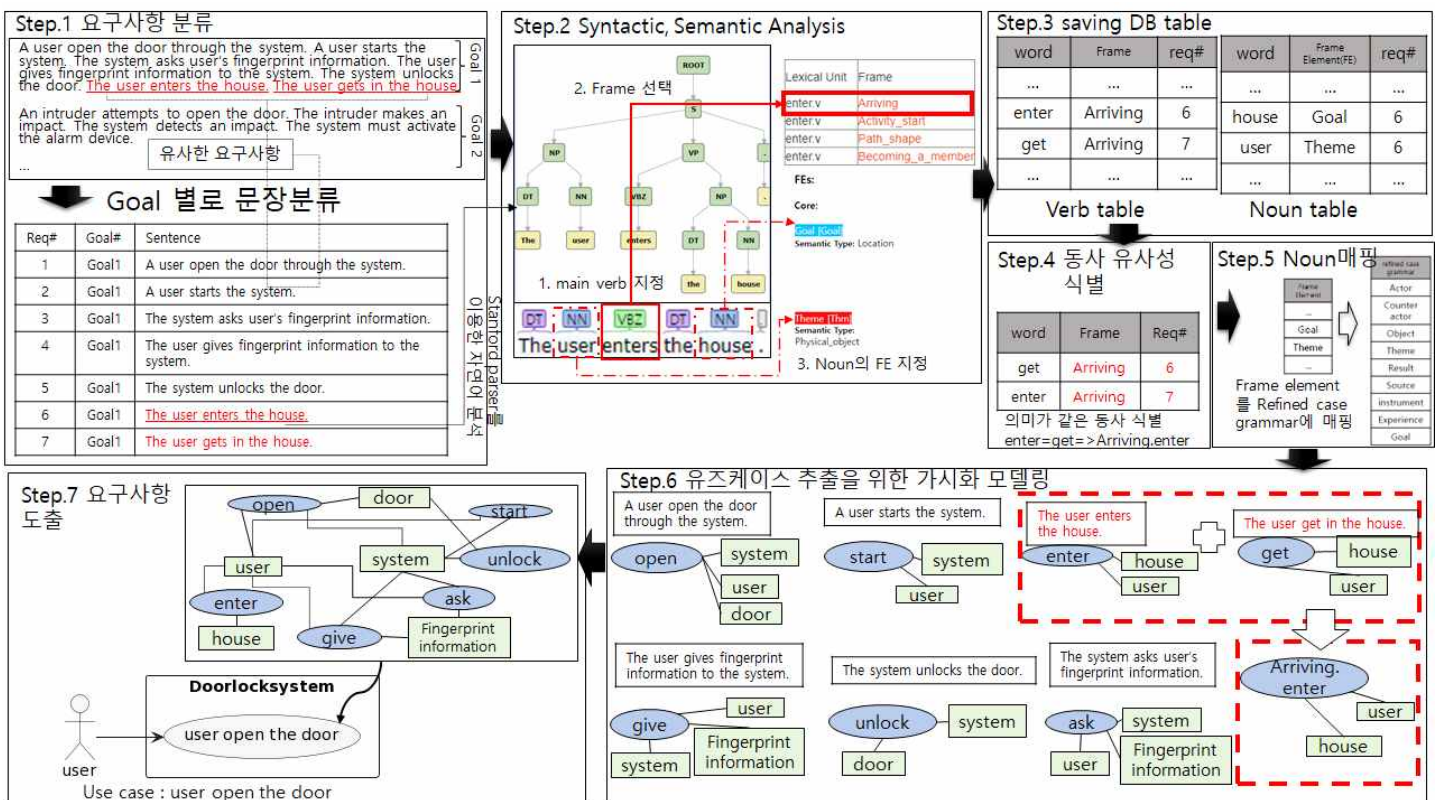
3. 동사 유사성 식별 기반 요구사항으로부터 유스케이스 생성 방법

본 논문은 1) Stanford parser를 통해 자연어 문장에서

구문론적 정보들을 식별하고, 2) 식별된 정보에서 품사(명사, 동사 등)를 구분하고, 3) FrameNet을 통해 명사, 동사들의 Frame 정보를 식별하여 DB에 저장한다. 4) 저장된 정보를 통해 유사성이 높은 동사를 식별하고, 5) 가시화 모델링을 통해 유스케이스를 생성한다.

그림 3은 동사 유사성 식별을 통한 요구사항으로부터 유스케이스 생성 방법이다. 구체적인 절차는 다음과 같다.

1. Step1은 비정형 자연어 요구사항을 문장 단위로 나눈다.
2. Step2는 나뉘어진 문장을 구문론적/의미론적 분석한다.
 - 2.1. Stanford parser를 통해 문장 내 품사를 태깅한다.
 - 2.2. 태깅된 품사 중 주요 동사를 식별한다.
 - 2.3. 주요 동사에 해당되는 Frame을 식별한다.
 - 2.4. 주요 동사 외의 품사들의 FE를 식별한다.
3. Step3은 앞서 식별된 정보들을 DB 테이블에 저장한다.
 - 3.1. Verb Table에 동사들의 정보(word, frame, req.)를 저장한다.
 - 3.2. Noun Table에 명사들의 정보(word, FE, req.)를 저장한다.
4. Step4는 주요 동사들의 유사성을 식별하고 단일화 한다.
 - 4.1. SQL 질의어를 사용하여 Verb Table에서 동일한 Frame을 갖는 주요 동사를 추출한다.
 - 4.2. 추출된 동사들에서 Req. 번호가 다른 동사만 다시 추출한다.
 - 4.3. 다시 추출된 동사들이 같은 의미로 사용된 단어인지 비교한다. 같은 의미로 사용된 동사는 단일화 한다.
5. Step5는 Noun Table에 저장된 명사들에 Refined Fillmore' s case grammar의 격(case)을 지정한다.
 - 5.1. 각 명사들에 해당되는 격(Actor, Counter actor,



- Object, Theme, Result, Source, Instrument, Experience, Goal)을 지정한다.
6. Step6은 유스케이스 추출을 위한 가시화 모델링을 수행한다.
 - 6.1. 각 문장의 주요 동사를 기준으로 명사를 실선으로 연결하여 모델화한다.
 - 6.2. 의미가 같은 동사들을 하나로 합친다.
 - 6.3. 동일한 명사들을 하나로 합친다.
 7. Step7은 합쳐진 모델을 통해 유스케이스 다이어그램을 생성한다.

4. 결론 및 향후 연구

본 논문은 비정형 요구사항에서 동사의 유사성을 식별하여 유사한 요구사항 문장을 합치는 방법을 제안한다. 이 방법은 기존 비정형 요구사항으로부터 UCP 기반 비용 예측 방법에서 해결하지 못했던 요구사항 내 유사한 동사들을 식별할 수 있다. 이를 통해 자연어에 등장하는 유의어를 하나의 유스케이스로 합쳐 줄일 수 있다. 그리고 기존 가시화 모델링을 통한 유스케이스 추출 방법을 개선하였다.

향후 연구로써, 명사의 유사성 측정 방법과 인공지능 기법을 이용한 자동화 프로세스를 연구할 예정이다.

Acknowledgment

이 논문은 교육부 및 한국연구재단의 4단계 두뇌한국21 사업의 지원(F21YY8102068)과 2022년도 정부(교육부)의 재원으로 한국연구재단의 지원(No. 2021R111A305040711, No. 2021R111A1A01044060)을 받아 수행된 연구임.

참고문헌

- [1] Fernández, D.M., Wagner, S., Kalinowski, M. et al. "Naming the pain in requirements engineering". Empirical Software Engineering 22, 2298–2338 (2017).
- [2] Gustavo Pinto, "Naming the Pain in Developing Scientific Software" IEEE SOFTWARE, 2020
- [3] 이현승, 이윤선 "공공SW 사업 발주관리의 현황, 문제점, 개선방안" 월간SW중심사회 No.41 2017
- [4] Josef Ruppenhofer et al "FrameNet II: Extended Theory and Practice"
- [5] Park, Bo Kyung, and R. Young Chul Kim. "Effort Estimation Approach through Extracting Use Cases via Informal Requirement Specifications" Applied Sciences 10, no. 9: 3044, 2022