강박스 교량을 대상으로 한 구조계산서의 누락된 설계항목 검토 자동화 방법론

A Method for Automatic Check of Omitted Design Item in Structural Calculation Document of Steel Box Bridges

박 상 일* • 안 현 정** • 김 봉 근*** • 이 상 호****
Park, Sang-Il • An, Hyun-Jung • Kim, Bong-Geun • Lee, Sang-Ho

ABSTRACT

A method for automatic check of omitted design item in structural calculation document of steel box bridges is proposed. A method for automatic check of omitted design item in structural calculation document of steel box bridges is proposed. Information processing for the proposed method is divided into two steps: automatic generation of document structure in XML Schema Definition (XSD) format and extract omitted design items by using the XML Schema matching technique. The automatic omitted element filter is developed on the basis of the proposed method, and the accuracy of the developed module is examined with case study subjected to existing structural calculation document samples.

Keywords: *Structural calculation document, steel box bridge, XML, schema matching*

1. 서 론

컴퓨터 및 인터넷의 급속한 발전으로 인해 정보는 손쉽게 다른 사람과 공유가 가능하게 되었고, 이런 정보들이 쌓여, 작금에 이르러서는 개인 한 사람이 그 정보를 조정 할 수 있는 양의 범위를 넘어서고 있다. 특히, 사회기반시설물의 유지관리 같은 부분에 있어서는 그 시설물의 정확한 정보를 적시적기에 제공하는 것이 유지관리를 하는데 있어서 가장 중요한 요소로 볼 수 있는데, 너무나도 많은 양의 정보는 이제는 필요한 정보를 찾는데도 많은 시간을 허비하게 하기도 한다. 특히 구조계산서와 같은 설계문서는 구조물을 설계한 조건과 해석한 결과 등 구조물의 세부적인 역학적 정보를 담고 있기 때문에 구조물의 유지관리 단계에서 발생하는 각종 업무에 기본적인 기준이 되는 반면에 너무 방대한 양의 정보를 담고 있기 때문에 세세한 항목의 누락 여부는 쉽게 찾아내기 힘들다. 최근 국내외에서는 교량설계자동화와 관련한 다양한 프로그램이 개발 및 상용화되고 있다. 이창근 등(1999)은 설계검토를 위한 자동화 시스템을 개발하였으며, 박태권 등(2002)은 도로설계 검사 시스템 개발을 통해 구조물의 구조적인 설계검토를 하였다. 또한 김병진 등(2004)은 제작, 시공

* 학생회원 • 연세대학교 사회환경시스템공학부 석사과정 Email: sang_il@csem.yonsei.ac.kr
** 연세대학교 사회환경시스템공학부 연구원 Email: ahj38@csem.yonsei.ac.kr
*** 정회원 • 연세대학교 사회환경시스템공학부 박사과정 Email: bgkim@csem.yonsei.ac.kr
**** 정회원 • 연세대학교 사회환경시스템공학부 정교수 Email: lee@yonsei.ac.kr

및 유지관리와 연계된 설계자동화 시스템 개발을 통하여 유지관리에 필요한 구조설계검토 방법을 제시한바 있으며, 최근 다양한 설계자동화 프로그램이 상용화되고 있다. 그러나 현재 데이터베이스화된 설계기준이 설계검토에 활용되는 과정은 프로그램화 되어 있는 절차에 따르는 것이며, 기 작성된 설계정보나 전통적으로 행해진 실무를 통해 작성된 설계정보를 검토하는 작업은 아직까지 인력에 의존하는 수동적 방법으로 수행되고 있다. 따라서 본 연구에서는 이진훈 등(2007)과 Lee 등(2006)이 개발한 방법론을 사용하여 스키마 매칭을 이용한 구조계산서 항목에서 누락 항목을 자동을 검토 추출해주는 방법에 관한 연구를 수행하였다.

2. 구조계산서의 누락항목 검토 방법

그림 1은 구조계산서 검토 시 누락 항목을 자동을 추출 해 내기위한 방법의 개략적인 흐름도를 나타낸 것으로, 크게 텍스트 파일 문서의 목차를 정의 및 XML 스키마를 생성하기 위한 부분과 표준 스키마 파일과의 스키마 매칭 부분, 그리고 누락항목 추출 부분으로 나눌 수 있다.

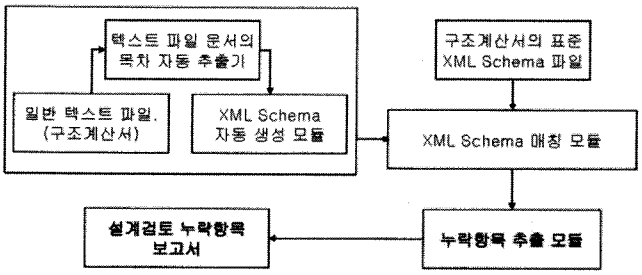

그림 1 설계 검토 시 누락 항목 자동 추출을 위한 흐름도

2.1 구조계산서의 자동 XML Schema file 생성

일반적으로 스프레드 쉬트 형태로 작성되어 있는 구조계산서의 스키마 매칭을 위해서는 그 문서의 항목들에 각각의 계층 정보를 담는 것이 필수적이다. XML Schema Definition(XSD)는 XML 문서에서의 각 요소를 형식적으로 기술하는 방법을 규정한 W3C의 권고사항이다. 이 XSD에서는 문서정보의 구조가 트리형식으로 특성화 되어 있다. 이에 이진훈 등(2007)은 문서의 각 머리글자를 상호 비교하여 그 계층정보를 문서에 주입하고, 그 정보에 따라 XSD 파일을 생성하는 연구를 하였다. 이렇게 생성된 XSD 파일은 표준 XML Schema 파일과의 스키마 매칭을 하여 그 누락항목을 추출해 낼 수 있다.

XSD 각 요소의 특성은 요소명, 부모요소, 자식요소, 형제요소로 구분할 수 있으며, 이를 각 요소의 특성집합으로 표현하면 식 1과 같다(김봉근 등, 2006).

$$D = \{D_N, D_P, D_B, D_C\} \tag{1}$$

여기서 D는 각 요소가 가지는 특성집합을 의미하며, D_N은 요소집합, D_P는 부모요소집합, D_B는 형제요소집합, 그리고 D_C는 자식요소 집합을 의미한다.

2.2. Similarity measure와 relaxation labeling 방법을 이용한 스키마 매칭

스키마 매칭은 스키마 정보의 조작처리에서 기본적인 연산이 서로 다른 두 스키마 사이에서 시맨틱 정보를 가지는 서로 상응하는 요소들을 대조하는 기법이다. 본 연구에서의 Tversky와 Shafir(2004), Yi 등(2005)이 제안하고, Lee 등(2006)이 구조계산서의 스키마 매칭을 위해 수정한 방법을 이용하였다. 이 스키마 매칭은 크게 두 가지 과정을 거치게 되는데, 첫 번째로 각 요소간의 언어적 유사성을 측정하는 과정을 거치게 되며, 이후 요소간의 계층적 구성요건을 제약조건으로 하는 relaxation labeling 과정이 수행된다.

식 2는 각 요소간의 언어적 유사도를 측정하기 위해 수행되는 similarity measure의 방법에 관한 식을 나

타낸 것으로 한 문서의 i번째 항목과 다른 문서의 j번째 요소의 요소명 집합 D^i와 D^j 사이의 특징별 유사도 $S(D^i, D^j)$들의 합을 나타낸 것이다. 여기서 δ_n, δ_p, δ_b, δ_c는 각각 요소명, 부모요소, 형제요소 및 자식요소에 대한 가중치를 의미하며 그 합은 1이 된다. 각각 k, l개의 범주 개체수를 가진 두 스키마의 XSD 파일이 주어졌을 때, 유사성 척도 행렬은 식 3과 같이 얻을 수 있다.

$$S_T(D^i, D^j) = \delta_n S_N(D_N^i, D_N^j) + \delta_p S_P(D_P^i, D_P^j) + \delta_b S_B(D_B^i, D_B^j) + \delta_c S_C(D_C^i, D_C^j) \tag{2}$$

$$S = \begin{bmatrix} S_T(1,1) \cdots S_T(1,l) \\ \vdots \quad \ddots \quad \vdots \\ S_T(k,1) \cdots S_T(k,l) \end{bmatrix} \tag{3}$$

여기서 $S_T(i,j)$는 서로 다른 두 스키마에 각각 존재하는 i요소와 j요소 사이의 유사도이고, 그 값은 두 요소간의 유사성을 나타내는 정량적 값이 된다.

유사도를 이용한 요소매칭은 두 스키마에서 그 단어의 유사성과 부모, 형제, 자식 요소의 특성까지만 고려하기 때문에 전체적인 구조적인 맥락을 고려할 수 없다. 그러므로 relaxation labeling의 방법을 적용하여 서로 다른 두 스키마가 가지는 계층적 구조의 특성을 고려한다. 식 4는 relaxation labeling의 방법을 적용하여 개선되는 신뢰도 식을 나타낸 것이다.

$$p_i^{(t+1)}(j) = \frac{p_i^{(t)}(j) q_i^{(t)}(j)}{\sum_{\beta=1}^{l} p_i^{(t)}(\beta) q_i^{(t)}(\beta)} \tag{4}$$

여기서 $p_i(k)$는 i요소가 j요소에 매칭이 될 수 있는 신뢰도를 나타낸 것이며, 그 초기 값으로는 similarity measure에서 나온 $S_T(D^i, D^j)$값을 사용하였다. t는 반복 횟수를 나타낸 것이며, l은 소스스키마의 요소의 수이고, q_i는 지지함수로 신뢰도의 지지도를 정량화하는 기능을 한다. 그리하여 반복계산에 의한 계산단계 t에서의 개선 값에 의한 요소매칭 신뢰도 행렬은 식 5와 같다.

$$P^t = \begin{bmatrix} P_1^t(1) \cdots P_1^t(l) \\ \vdots \quad \ddots \quad \vdots \\ P_k^t(1) \cdots P_k^t(l) \end{bmatrix} \tag{5}$$

2.3. 설계검토 누락항목 추출의 방법

구조계산서의 목차를 추출하여 XSD 파일로 변환한 파일은 표준화된 스키마 파일과 스키마 매칭을 통해 나온 그 결과 값에 따라 설계검토 누락항목을 판단하게 된다. 실무에서 작성한 구조계산서는 그 작성 기관이나 사람에 따라서 작성 형식이나 순서, 글자의 표현방법이 다른 경우가 대부분이다. 작성 형식, 순서 등은 앞 절에서 설명한 목차자동 추출기나 XML Schema에 의해 처리가 가능하다. 그러나 같은 항목이지만 그 표현방법이 한글과 영문이 혼용되어 있는 경우 혹은 유사의미의 단어를 사용하는 경우에는 단어별 매칭의 정확도가 확연히 떨어진다. 이에 본 연구에서는 구조계산서에서 주로 사용하는 영문 및 유사 단어를 데이터베이스로 저장을 시켜 자동으로 표준 항목으로 바꾸게 하였고, 또한 항목을 단어가 아닌 단어의 글자별로 그 유사도를 비교하는 방법을 선택하여 연구를 수행하였다. 그리하여 스키마 매칭 결과 값 중 1의 항목이 없으면 그 항목은 누락되었다고 판단하였다.

3. 설계 검토 누락항목 자동 추출 모듈의 정확성 검증

본 연구에서 설계한 모듈의 정확성을 검증하기 위해 우선 실제 구조계산서의 일부분으로 테스트를 해보았다. 그림 2는 표준 스키마와 구조계산서의 XSD 파일의 매칭 경로를 나타낸 것이며, 표 1은 표준화된 스키마 모델과 구조계산서의 보강재의 설계 항목에 대해 스키마 매칭을 통해 나온 신뢰도 값을 나타낸 것이다. 표준화된 스키마에서 보강재의 설계 부분에서의 항목은 총 17개이고, 구조계산서에서 보강재의 설계부분 항목은 총 9개이다.

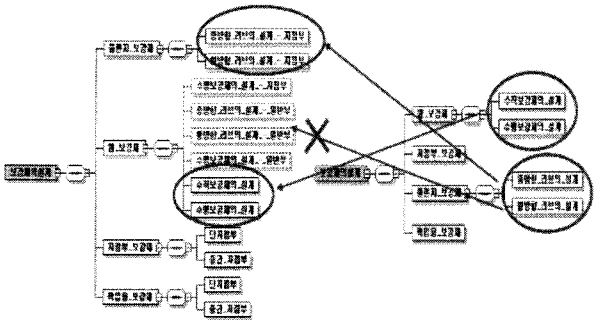

그림 2 표준 스키마(좌)와 실제 구조계산서(우)의 보강재의 설계부분에 관한 XSD 파일

표 1에서 보는 바와 같이 '보강재의 설계' 등 항목의 이름이 같은 부분은 그 신뢰도가 '1'이 나와 완전히 일치 하는 것을 볼 수가 있다. 하지만 '종방향 리브의 설계'와 '횡방향 리브의 설계'는 일반부와 지점부 각각 두개씩의 항목이 있지만, 지점부 부분에만 매칭이 된다. 그 이유는 그 부모요소가 '플랜지 보강재'로 같기 때문에 상대적으로 그 유사도가 높기 때문이다. 그리하여 실제 구조계산서의 예에서 '수평보강재의 설계 - 지점부', '종방향 리브의 설계 - 일반부', '횡방향 리브의 설계 - 일반부', '수평보강재의 설계 - 일반부', '단지점부' 2개, '중간 지점부' 2개는 매칭이 되지 않는 항목으로 표준화된 스키마와 비교하여 누락된 항목이라고 할 수 있다.

이와 같은 방법으로 실제 구조계산서의 목차를 자동 추출하여 계층을 정의하고 XSD 파일을 생성한 후 표

표 1 두 개의 XSD 파일을 스키마를 매칭 시킨 후의 그 신뢰도 값

표준 / 실제	보강재의 설계	플랜지 보강재	종방향 리브의 설계(지점부)	횡방향 리브의 설계(지점부)	웹 보강재	수평보강재의 설계 (지점부)	종방향 리브의 설계(일반부)	횡방향 리브의 설계(일반부)	수평보강재의 설계 (일반부)
보강재의설계	1	0	0	0	0	0	0	0	0
웹 보강재	0	0	0	0	1	0	0	0	0
수직보강재의 설계	0	0	0	0	0	0	0	0	0
수평보강재의 설계	0	0	0	0	0	0	0	0	0
지점부 보강재	0	0	0	0	0	0	0	0	0
플랜지 보강재	0	1	0	0	0	0	0	0	0
종방향 리브의 설계	0	0	1	0	0	0	0	0	0
횡방향 리브의 설계	0	0	0	1	0	0	0	0	0
작업용 보강재	0	0	0	0	0	0	0	0	0

표준 / 실제	수직보강재의 설계	수평보강재의 설계	지점부 보강재	단지점부	중간 지점부	작업용 보강재	단지점부	중간 지점부	
보강재의설계	0	0	0	0	0	0	0	0	
웹 보강재	0	0	0	0	0	0	0	0	
수직보강재의 설계	1	0	0	0	0	0	0	0	
수평보강재의 설계	0	1	0	0	0	0	0	0	
지점부 보강재	0	0	1	0	0	0	0	0	
플랜지 보강재	0	0	0	0	0	0	0	0	
종방향 리브의 설계	0	0	0	0	0	0	0	0	
횡방향 리브의 설계	0	0	0	0	0	0	0	0	
작업용 보강재	0	0	0	0	0	1	0	0	

준 스키마와 스키마 매칭을 수행하였다. 스키마 매칭은 3가지의 경우로 나누어 수행하여 그 정확성을 분석하였으며, 3가지의 경우는 각각 전체 스키마를 한 번에 비교하는 경우, 구조계산서 root 요소이하 첫 번째 단계에 출현되는 요소를 기준으로 스키마를 분리하여 비교하는 경우 그리고 마지막으로 첫 번째 단계에 출현되는 요소 중 하위 요소에 동일한 의미를 가지는 요소명이 중복적으로 다수 출현되는 요소를 한 단계도 세분화하여 비교하는 경우이다. 스키마 매칭 대상이 되는 표준 스키마의 항목 수는 총 313개였으며, 실제 구조계산서의 항목 수는 총 277개였다.

첫 번째로 전체 스키마를 한 번에 비교한 경우 자동추출을 잘못한 항목수가 61개이어서 스키마 매칭의 정확도는 약 80.51%를 나타내었다. 이는 구조계산서의 경우 매우 많은 항목이 결합되어 XSD 파일의 형태가 복잡하기 때문에 다른 부모의 요소에 속한 요소 중에 같은 글자가 많이 존재 하며, 이로 인하여 다른 위치의 글자끼리의 유사도가 매우 높게 측정됨으로써 전체적인 스키마 매칭에 있어 정확도가 떨어진다고 판단된다. 두 번째로 root 요소이하 첫 번째 단계 요소를 기준으로 세분화하여 스키마 매칭을 수행한 는 표 2에 나타내었다. 표 2에서 알 수 있듯이 표준 스키마와 실제 구조계산서를 설계기준, 바닥판, 주형설계, 이음부 설계, 보강재설계, 세로보 및 가로보, 다이아프램, 전단연결재, 사용성검토, 솟음, 신축이음설계 부분으로 나누어 각각의 항목에 맞게 스키마 매칭을 실행하였으며, 그 평균적인 정확도는 약 83.07%로 하나의 파일로 스키마 매칭을 실행하였을 때보다 정확도가 향상되었다. 마지막 세 번째 스키마 매칭 실험의 경우 그 추출 정확도가 100%가 아닌 바닥판과 주형설계, 세로보 및 가로보 항목을 더 세분화 하여 스키마 매칭을 실시하였고, 그 결과 누락 항목 추출의 정확도는 주형 설계부분 이외에는 모두 100%로 나왔다. 즉, XSD 파일이 간단할수록 그 정확도는 훨씬 높아졌다. 하지만, 기존에 작성된 구조계산서에서는 항목별로 분류를 하여 파일을 저장해

표 2 1차 분리에 의한 누락 항목 추출 결과

	Ⅰ	Ⅱ	Ⅲ	Ⅳ
설계 기준	31	30	0	100
바닥판	40	53	8	80
주형 설계	84	69	28	66.67
이음부 설계	32	11	0	100
보강재 설계	17	9	0	100
세로보 및 가로보	50	48	17	66
다이아프램	26	26	0	100
전단 연결재	12	12	0	100
사용성 검토 / 솟음 / 신축 이음 설계	21	19	0	100
계	313	277	53	83.07

표 3 2차 분리에 의한 누락 항목 추출 결과

	Ⅴ	Ⅰ	Ⅱ	Ⅲ	Ⅳ
바닥판	캔틸레버부(좌)	24	20	0	100
	캔틸레버부(우)	24	20	0	100
	중앙 지간부	15	13	0	100
	내부지점부 바닥판의 주장율 검토	1	0	0	-
계		64	53	0	100
주형설계	플랜지 및 바닥판의 유효폭 계산	3	3	0	100
	작용하중의 산정	27	24	4	85.19
	스프라이스 설계	23	19	0	100
	뒴비틀림에 대한 검토	10	7	0	100
	단면 검토	21	16	4	80.95
계		84	69	8	90.48
세로보및가로보	세로보의 설계	18	10	0	100
	가로보의 설계	24	10	0	100
	가로보의 연결	8	14	0	100
	세로보의 연결	8	14	0	100
계		58	48	0	100

Ⅰ: 표준스키마의 항목 수, Ⅱ: 실제 구조계산서의 항목 수, Ⅲ: 잘못 추출한 항목 수,
Ⅳ: 정확도(%), Ⅴ: 표준스키마의 세부항목의 항목 수

높은 경우가 많지 않기 때문에 하나의 구조계산서를 적절한 항목으로 나누어 스키마 매칭 수행한다면 보다 정확도를 향상시킬 수 있을 것으로 판단된다.

4. 결 론

본 연구에서는 기존에 작성되어 있던 전자 문서 형식의 구조계산서의 설계 검토 누락 항목을 자동으로 추출하는 방법론을 제시하였고, 그 방법론에 따라 실제 구조계산서에 적용 가능한 응용모듈을 개발하였으며, 이에 따라 실제 구조계산서를 대상으로 그 모듈의 정확성을 검증해 보았다. 표준 스키마의 313개의 항목과 실제 구조계산서 항목 277개의 스키마 매칭의 결과 그 누락 항목 추출 정확도는 약 80.51%로 비교적 높은 수치를 보였으며, 더욱이 스키마 파일을 더 세분화하여 스키마 매칭을 시킨 결과는 그 정확도가 더욱 높게 나타나 누락 항목 거의 대부분을 추출해 내었다. 이에 따라 전체 문서상의 적절한 위치에서 문서를 분리한 후 스키마 매칭의 방법을 이용하여 설계 검토 누락 항목을 추출하면, 이로 인해 설계 검토의 업무 정확성 및 효율성이 크게 증대될 뿐만 아니라, 시설물의 유지관리에 관련한 엔지니어들에게 질 높은 정보의 제공이 가능하여 긴급 상황 등이 발생하였을 때 더욱 신속하게 대처할 수 있을 것으로 기대된다.

감사의 글

본 연구는 건설교통부에서 실시한 건설핵심기술연구개발사업(교량설계핵심기술연구단)의 연구비 지원에 의해 수행되었으며, 2007년도 교육인적자원부 BK21사업의 일환인 연세대학교 사회환경시스템공학부 미래사회기반시설 산학연공동사업단의 부분적인 지원을 받았음.

참고문헌

김병진 외 22명 (2004) 제작, 시공 및 유지관리와 연계된 강교 설계자동화 시스템 개발, R/D /02 산학연 A01-06, 한국건설교통기술평가원

김봉근, 정연석, 김동현, 이상호 (2006) XML 스키마 매칭 기법을 이용한 구조설계 문서구조 표준화방법론, 한국전산구조공학회 2006 정기 학술대회 논문집, 한국전산구조공학회, 19(1), pp.200~207.

박태권 외 13명 (2002) 도로설계 검사 시스템 개발 (설계검사 및 안전도 평가모듈개발), 건설교통부

이진훈, 김봉근, 박상일, 이상호 (2007) 사회기반시설물 안전관리 지원 데이터베이스 구축을 위한 구조계산서의 XML 문서화 방법론, 2007년도 한국방재학회 학술발표대회 논문집, pp.414~417.

이창근, 이원태, 서영국, 허노윤 (1999) 강박스교 설계검토 자동화 시스템 개발, 한국도로공사 도로연구소

Lee, Sang-Ho, Kim, Bong-Geun, Kim, Dong-Hyun and Jeong, Yeon-Suk (2006) Development of standardized semantic model for structural calculation documents of bridges and XML Schema matching technique, *Proceedings of the Third International Conference on Bridge Maintenance Safety and Management (IABMAS)*, pp.633~634, CD Rom paper P-228 (8 pages), Porto, Portugal.

Tversky, A. and Shafir, E. (2004) *Preference, Belief, and Similarity*, A Bradford Book, The MIT Press Cambridge.

Yi, Shanzhen, Huang, Bo, and Chan, Weng Tat (2005) XML application schema matching using similarity measure and relaxation labeling, *Information Sciences*, 169(1), pp.27~46.