

# 사회기반시설물 안전관리 지원 데이터베이스 구축을 위한

## 구조계산서의 XML 문서화 방법론

### A methodology for XML documentation of the structural calculation document to build database supporting safety management of infrastructures

이진훈\* · 박상일\*\* · 김봉근\*\*\* · 이상호\*\*\*\*

Kim, Bong Geun · Park, Sang Il · Lee, Jin Hoon · Lee Sang Ho

#### Abstract

A methodology for XML documentation of the structural calculation document is presented to support manipulation of the design information on the internet. The text file format is chosen as a neutral format because it can be easily translated from office documents generated from engineering practice. The first word of each line is compared with the reserved numbering groups, and relative levels among the lines are defined to generate the hierarchically structured XML document of the text file. The demonstration subjected to sample general documents and structural calculation documents shows that the prototype application module based on the developed methodology can be adopted to build the database of design information which supports the safety management of infrastructures.

**key words** : Structural calculation document, XML documentation, Database, Safety management of infrastructures

#### 1. 서 론

사회기반시설물의 안전관리 업무를 효율적으로 수행하기 위해서는 그 시설물에 대한 정확한 정보를 적시 적기에 엔지니어에게 제공할 수 있는 정보체계 구축이 필요하며, 이는 곧 시설물의 안전관리를 위한 정보시스템 구축의 가장 기본적인 목적이라 할 수 있다. 구조계산서는 설계조건, 구조해석 결과, 안전성 및 사용성 검토결과와 같은 해당 구조물의 설계에 따른 다양한 정보를 포함하고 있기 때문에 구조물의 안전점검 및 안전진단 업무에서 참조할 수 있는 가장 기본적인 설계문서의 하나이다. 그러나 건설교통부(2004)와 한국시설안전기술공단(2004)과 같은 시설물의 유지관리를 담당하고 있는 주요 공공기관들은 도면의 경우 재활용 가능한 형태의 파일 포맷을 납품받고 있으나 구조계산서와 같은 문서의 경우 이미지 형태의 문서를 납품 포맷으로 규정하고 있기 때문에 구조계산서에 포함된 정보를 재가공하여 엔지니어에게 적시적기에 제공할 수 없다.

XML(eXtensible Markup Language)은 정보를 계층적으로 구조화할 수 있으며, 정보의 구성요소를 사용자가 정의할 수 있는 그 특유의 확장성으로 인하여 다양한 분야에서 정보 표준화를 위한 기반 표준으로 채택되고 있다. 한국건설기술연구원(2004)에서는 국내 건설분야에서 사용되는 다양한 문서를 조사·분석하여 전자문서별 정보체계를 XML을 기반으로 개발하였다. 그러나 아직 구조계산서와 같은 엔지니어링 문서의 경우 의미적 요소가 충분히 반영되어 있지 못하다. 이와 같은 배경에 따라 양영애 등(2005)은 강교량 구조계산서를 대상으로 한 표준 XML Schema에 대한 연구를 수행한 바 있으며, 김봉근 등(2006)은 이러한 엔지니어링 문서의 표준 XML Schema를 효율적으로 개발하기 위한 방법론을 제시하였다. 그리고 이상호 등(2006)은 교

\* 학생회원 · 연세대학교 사회환경시스템공학부 토목환경공학전공 · 석사과정 · E-mail: jeffyi@csem.yonsei.ac.kr

\*\* 학생회원 · 연세대학교 사회환경시스템공학부 토목환경공학전공 · 학부생

\*\*\* 정회원 · 연세대학교 사회환경시스템공학부 토목환경공학전공 · 박사과정

\*\*\*\* 정회원 · 연세대학교 사회환경시스템공학부 토목환경공학전공 · 정교수

량 구조계산서의 표준 XML Schema를 이용하여 구조계산서 정보를 데이터베이스화 하고 이를 STEP 기반의 3차원 형상정보와 연계한 시범 시스템을 개발한 바 있다. 그러나 아직 실무에서 XML 형태로 문서를 작성하는 것이 용이하지 않기 때문에 보다 실용성이 있는 XML 문서화를 위한 연구가 필요하다. 따라서 본 연구에서는 실무에서 작성된 문서의 텍스트 파일 포맷을 읽어서 그 계층 정보를 자동으로 인식하여 구조계산서의 문서정보를 XML 문서로 변환하는 방법에 관한 연구를 수행하였다.

## 2. 구조계산서 문서항목의 자동 계층 분류 방법

본 연구에서는 구조계산서의 XML 문서화를 위해 구조계산서 상의 제목들을 읽어 들여 그 문장의 맨 처음에 있는 머리 글자와 미리 저장된 머리 글자의 그룹을 비교하여 상호간의 계층적 레벨을 분류하는 방법을 개발하였다. 그림 1은 이와 같은 방법의 전체적인 흐름도를 나타낸 것으로서 그림 1에 나타낸 바와 같이 항목의 구분을 자동화하기 위한 절차는 크게 두 단계로 구성될 수 있다. 첫 번째 단계는 텍스트 문서내의 각 행이 단순 내용인지 혹은 목차인지를 결정하는 단계이고, 두 번째 단계는 첫 번째 단계에서 목차로 분류된 각 항목간의 상·하위 관계를 정의하는 것이다.

첫 번째 단계에서는 텍스트 문서에 저장되어 있는 각 정보를 줄마다 읽어서 제일 첫 번째 단어만 따로 저장하고 그 문자를 미리 입력해 둔 데이터와 비교를 하여 목차항목으로 구분된 경우 별도의 식별번호를 부여한다. 미리 입력해둔 데이터란 머리 글자로 나올 수 있는 단어들을 모아둔 것을 의미한다. 이는 숫자나 기호는 다르지만 같은 위상에 나타날 수 있는 머리 글자들을 분류하기 위해 정의하였으며, 본 연구에서는 다양한 형태의 구조계산서를 수집하여 그 데이터베이스를 구축하였다. 만약 구축되어 있는 데이터베이스와 일치하는 머리 글자가 없으면 그것은 목차가 아닌 단순 내용을 기술한 부분으로 가정하였다.

두 번째 단계에서는 각 그룹별로 식별번호가 정의된 머리 글자들을 그 앞 또는 뒷줄과의 머리 글자의 식별번호와 상호 비교하여 그 상대적인 계층관계를 설정한다. 문서의 맨 처음 읽어 들인 줄은 가장 상위 등급으로 분류를 하였으며, 그 다음 줄을 읽어서 뒷줄과 같은 범주가 아니라면, 그 줄은 뒷줄보다 한 단계 하위 등급으로 구분을 하였다. 이런 형식으로 문서내의 각 행을 처리해 나가다가 뒷줄과 같은 범주의 머리 글자가 나오게 된다면, 그 줄은 같은 범주에 속하는 줄과 같은 등급의 레벨로 정의하였다. 이때에는 그 줄보다 하위

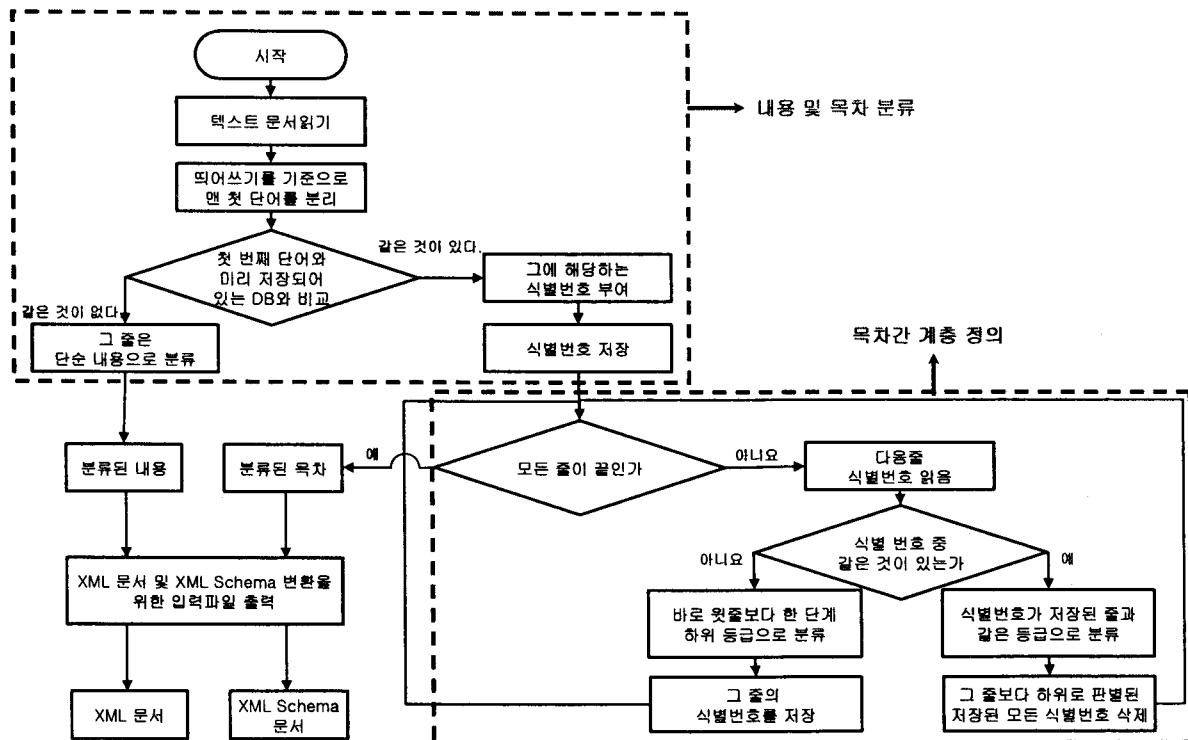


그림 1. 구조계산서 문서구조의 자동 계층분류 흐름도

등급으로 분류된 기존에 저장되어 있던 모든 식별번호를 삭제해 주어야 한다. 이러한 이유는 구조계산서와 같이 매우 복잡하고 상세분류까지 정의된 문서의 경우 위에 상위에서 식별된 줄들과는 다른 순서로 머리 기호가 사용되는 사례가 매우 많기 때문이다.

이와 같은 과정에 따라 정의된 각 목차의 제목과 내용은 김봉근 등(2006)이 XML 문서와 XML Schema를 생성하기 위해 미리 정의한 양식에 따라 출력되고, 이를 이용하여 XML 문서와 XML Schema 파일을 동시에 생성할 수 있다. 단순히 XML 문서를 생성하지 않고 해당 문서구조를 표현하는 XML Schema를 생성하는 이유는 스키마 매칭 기법을 이용하여 읽어들이 구조계산서 파일을 표준화된 XML Schema에 따른 구조로 변환하기 위함이다.

### 3. 구조계산서의 XML 문서 변환 사례 및 정확성 검토

본 연구에서는 2장에서 설명한 구조계산서 문서 구조의 자동 계층 분류 방법론에 따른 응용모듈을 개발하였으며, 그림 2는 수집된 구조계산서를 대상으로 변환한 사례를 나타낸 것이다. 그림 2에 나타난 바와 같이 각 제목의 머리 글자를 기준으로 계층화한 정보는 그 상·하위를 구분하는 기호에 따라서 '[A]', '[B]' 등과 같은 미리 정의된 양식으로 자동으로 태그가 생성되고, 이 양식을 이용하여 구조계산서의 문서정보를 XML 문서와 XML Schema 문서로 변환한다.

개발된 응용모듈의 검증은 위해 먼저 보고서, 제안서, 논문과 같이 일반적으로 작성되는 다양한 문서를 대상으로 실험하였다. 실험결과 14개 실험문서의 총 1,173개 목차 항목 중 올바르게 구분하지 못한 항목은 14개로서 약 98.8%의 정확도를 보였다. 올바르게 구분하지 못한 경우는 모두 문서 작성자의 항목구분 작성 오류에 의한 것으로서 문서 작성자의 실수가 없는 문서의 경우 100%의 정확도를 나타내었다. 또한 본 연구에서 대

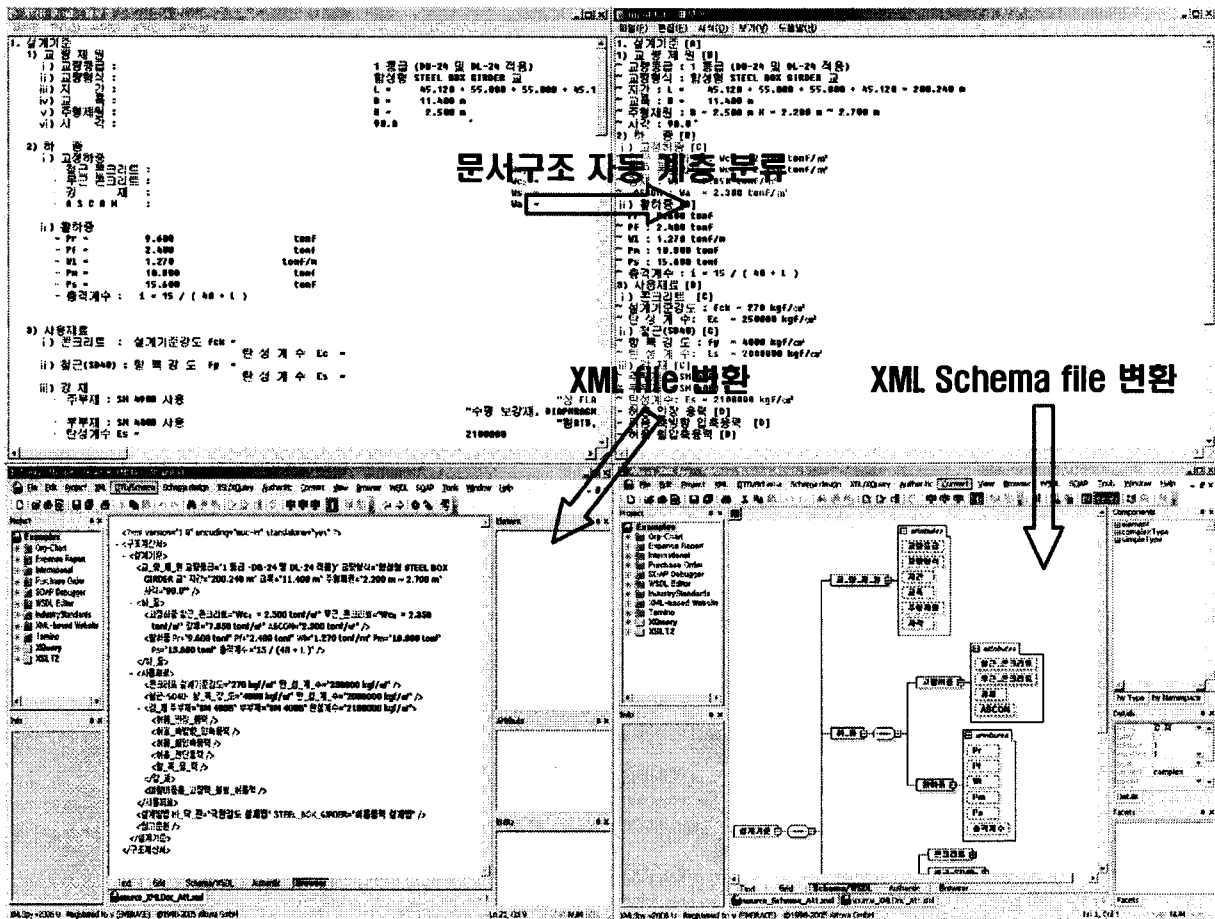


그림 2. 가덕IC교의 구조계산서를 XML file로 변환하는 과정

상으로 하는 구조계산서를 수집하여 동일한 방법으로 실험하였으며, 실험결과 35개 실험문서의 총 28,514개 목차항목 중 올바르게 구분하지 못한 항목은 1,145개로서 약 95.98%의 정확도를 보였다. 구조계산서를 대상으로 한 실험에서 발생된 오류는 의미적으로는 계층 구분이 다르지만 동일한 기호를 중복적으로 사용한 것에 기인한 것이다. 이러한 경우에는 머리 기호만으로 그 항목간의 계층을 분류할 수 없다. 이와 같이 현재 실무에서 작성된 구조계산서의 경우 정형화된 항목구분 형식이 없이 일부 비합리적인 항목구분 형식이 사용되기도 한다. 그러나 건설CALS/EC 단체표준(한국건설기술연구원, 2004)에서 제시된 항목구분 형식과 같은 표준화된 문서양식의 사용이 확산된다면 일반문서를 대상으로 한 실험에서도 알 수 있듯이 본 연구에서 제시한 방법론은 실무에서 작성한 구조계산서에 포함된 세부적인 설계정보를 XML 문서로 변환하는데 효과적으로 활용할 수 있을 것으로 판단된다.

#### 4. 결 론

본 연구에서는 기존에 작성되어 있던 구조계산서 파일내의 목차 항목을 자동으로 계층화하고, 계층화된 각 목차 항목을 의미적 요소로 정의한 XML 문서로 변환하는 방법론을 제시하였다. 또한 제시된 방법론에 따른 응용모듈을 개발하였으며, 보고서와 같은 일반적으로 작성된 문서와 실무에서 작성된 구조계산서를 대상으로 그 정확성을 검증하였다. 14개의 일반문서와 35개의 구조계산서를 대상으로 실험한 결과 구조계산서의 경우 일반문서에 비하여 상대적으로 낮은 정확도를 보였다. 그러나 일반문서를 대상으로 한 실험에서는 문서 작성자에 의한 오류가 없는 경우 100%의 정확성을 나타내었으며, 이에 따라 추후 구조계산서를 작성하는데 있어서 표준이 되는 양식의 배포와 그것의 사용이 확산된다면 본 연구에서 제시된 방법론은 실무에서 작성된 구조계산서의 세부 항목구분을 성공적으로 수행하고 이를 XML 문서로 변환할 수 있는 것으로 판단된다. 이와 같이 변환된 구조계산서의 XML 문서는 사회기반시설물의 유지관리 업무 지원을 위한 설계정보체계의 데이터베이스 구축에 효과적으로 사용될 수 있으며, 안전점검 및 안전진단과 같은 사회기반시설물의 안전관리 업무를 수행함에 있어서 구조물의 안전성 판단의 기본적인 기준이 될 수 있는 설계정보를 엔지니어가 사용하는 응용프로그램에 맞추어 효과적으로 제공하는데 기여할 수 있을 것으로 기대된다.

#### 감사의 글

본 연구는 건설교통부에서 실시한 건설핵심기술연구개발사업(교량설계핵심기술연구단)의 연구비 지원을 받아 연구되었으며, 이에 깊은 감사를 드립니다.

#### 참고문헌

1. 건설교통부 (2004) 도로 및 하천공사의 설계·준공도서 전자납품 편람 (V.1.0). 건설교통부.
2. 김봉근, 정연석, 김동현, 이상호 (2006) XML 스키마 매칭 기법을 이용한 구조설계 문서구조 표준화방법론. 한국전산구조공학회 2006 정기 학술대회 논문집, 한국전산구조공학회, 제19권, 제1호, pp. 200-207.
3. 양영애, 김봉근, 이상호 (2005) 강박스교 구조계산서 XML 시맨틱 모델의 스키마 매칭 기법적용. 2005년 한국전산구조공학회 봄 학술발표회 논문집, 한국전산구조공학회, 제18권, 제1호, pp. 680-687.
4. 한국건설기술연구원 (2004) 건설CALS/EC 전자문서 표준 (V.1.0). 한국건설기술연구원.
5. 한국시설안전기술공단 (2004) 설계도서 등의 사본작성 및 관리지침. 한국시설안전기술공단.
6. 이상호, 정연석, 김봉근 (2006) 교량 유지관리 지원을 위한 CAD/CAE 정보와 엔지니어링 문서정보의 통합 데이터베이스. 한국CAD/CAM학회논문집, 한국CAD/CAM학회, 제11권, 제3호, pp. 183-196.