

# MPEG-2 TS 형태의 테스트 데이터 자동 생성기의 구현

염선화<sup>○</sup> 최병주 박기웅\*

이화여자대학교 컴퓨터학과, \*삼성 전자

{992COG15, bjchoj}@mm.ewha.ac.kr, kwpark@swc.sec.samsung.co.kr

## Implementation of Testdata Generation Tool for MPEG-2 TS

Sunhwa Yeum<sup>○</sup> Byoungju Choi Kiwoong Park\*

Dept. of Computer Science & Engineering, Ewha Womans University

\*SAMSUNG ELECTRONICS CO. LTD

### 요 약

멀티미디어 표준인 MPEG-2 TS 형태의 테스트 데이터를 자동 생성하는 도구인 TDGT(Test Data Generation Tool)를 설계하고 구현하였다. TDGT는 테스트로부터 기본적인 입력사항을 받아들이고 이를 기반으로 테스트 데이터를 자동 생성한 후 결과물로 테스트 데이터 스크립트 파일과 테스트 데이터 시나리오 파일, 테스트 데이터 커버리지 분석 결과를 테스트자에게 제공한다. TDGT는 복잡하고 방대한 테스트 데이터를 생성하는데 요구되는 노력, 시간을 상당히 줄여주며 전문적인 지식이 없는 이도 쉽게 테스트 데이터를 생성할 수 있게 한다. 본 논문은 TDGT가 테스트 데이터를 자동 생성하는 구조와 TDGT의 분석을 제시하였으며 이 시스템의 prototype을 기술한다.

### 1. 서론

최근의 상용 시스템은 대부분 복잡한 구조의 데이터를 대량으로 처리한다. 이러한 특징으로 인하여 상용 시스템에는 다양한 예제가 존재할 가능성이 높으므로 상용화 되기 전에 철저한 테스트가 요구된다. 그러나 복잡한 데이터를 받아들여 처리하는 시스템을 테스트 하기 위해서 생성하는 테스트 데이터는 그 조합이 매우 많기에 모든 데이터의 조합을 테스트 하는 것은 거의 불가능하다. 따라서 모든 데이터의 조합 중에서 의미 있는 것만을 선별하는 작업이 요구된다. 이 경우 테스트 데이터를 수작업으로 생성하는 데에는 신뢰도 문제, 시간과 비용 문제 등 많은 무리가 따른다. 또한 데이터가 복잡한 구조를 가지고 있기에 처리되는 데이터에 대한 전문적인 지식이 없는 이가 테스트 데이터를 생성하는 것은 무리가 있다. 그러므로 방대하고 복잡한 데이터를 처리하는 시스템을 테스트 하기 위해서는 테스트 데이터를 자동으로 생성하는 도구의 개발이 요구된다.

처리하는 데이터가 복잡하고 방대한 시스템의 대표적인 예가 멀티미디어를 처리하는 시스템이다. 멀티미디어는 음성, 데이터, 영상을 통합적으로 취급하는 말이다[1]. 이러한 멀티미디어 처리 기술의 국제 표준은 MPEG(Moving Picture Expert Group)이다. 특히 MPEG-2 표준은 더욱 높은 비트율에서 고품질을 실현하기 위한 것으로 미디어 통합을 목적으로 하고 있는 국제 표준이며 방송, 가전, 통신 등의 모든 산업 분야에서 주목 받고 있다. 따라서 MPEG-2 표준을 대상으로 하고 있는 테스트 데이터 자동 생성기를 개발하는 것은 의미 있다.

본 논문에서는 디지털 TV 방송을 위한 MPEG-2 TS(Transport Stream)를 처리하는 DASE(Digital TV Application Software Environment) 시스템의 테스트 데이터를 자동 생성하는 도구인 TDGT(Test Data Generation Tool)를 설계하고 구현하였다. TDGT는 기본적인 사용자 입력을 받아들여 사용자가 원하는 종류의 테스트 데이터를 데이터 베이스를 참조하며 자동 생성한다. 그리고 사용자에게 테스트 데이터 스크립트 파일과 테스트 데이터 생성 시나리오, 테스트 데이터 커버리지 분석 결과 등을 제공한다.

본 논문의 구성은 다음과 같다. 2장에서는 자동 생성과 관련된 TDGT의 구조를 기술하며, 3장에서는 TDGT에 대한 분석을 기술한다. 4장에서는 구현된 TDGT의 prototype을 기술하고, 5장에서 결론 및 향후 연구 과제를 제시한다.

### 2. 테스트 데이터 자동 생성과 관련된 TDGT의 구조

TDGT는 DTV 방송을 위한 MPEG-2 TS 데이터를 자동 생성한다. 따라서 테스트 데이터를 생성하기 위해서 방송을 위한 시나리오를 계획하고 이를 MPEG-2 TS 표준에 맞게 구성하여야 한다. MPEG-2 TS 표준은 여러 개의 테이블 구조의 데이터로 구성된다. 각 테이블은 여러 개의 필드로 구성되어 있으며 각 필드는 값을 취하는데 있어서 각각의 규칙과 다른 필드와의 연관관계를 가지고 있다[2,3,4,5,6]. 따라서 본 논문에서 구현한 TDGT는 테스트 데이터를 자동 생성하기 위하여 방송 시나리오와 테스트 데이터의 구조, 값을 자동 생성한다. 전체 TDGT의 구조는 그림 1과 같으며 자동 생성과 관련된 부분은 테스트 데이터 생성부(그림 1의 ①)와 데이터 베이스(그림 1의 ②)이다.

이 장에서는 TDGT의 테스트 데이터 자동 생성과 관련된 구조를 기술하겠다.

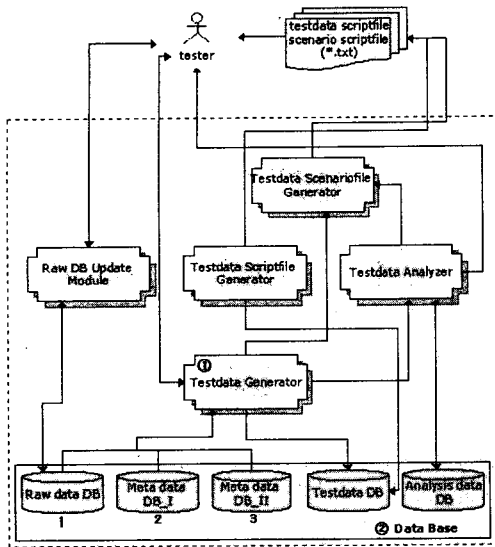


그림 1. TDGT의 전체 구조

2.1 방송 시나리오의 자동 생성

DTV 방송용 MPEG-2 TS는 단순한 유효 값들의 조합이 아니라 일정 시간 동안의 방송 프로그램의 계획을 담고 있어야 한다. 테스트 데이터를 자동 생성 하기 위해서 TDGT는 테스터에게서 요구사항을 입력 받아 이를 바탕으로 방송을 위한 시나리오를 계획한다. 방송 시나리오는 크게 채널 정보와 DTV 방송 프로그램 정보로 구성된다. 이때 채널 정보와 방송 서비스를 위한 어플리케이션 정보는 MPEG-2 표준의 규칙이나 제약 사항 등을 통해 자동 생성 할 수 없는 정보이다. 따라서 TDGT는 이들 정보를 Raw data DB(그림 1의 ②-1)로 구축하고 방송 시나리오를 계획 할 때 이 데이터 베이스에서 테스터의 요구 사항에 부합하는 데이터를 선택한다.

2.2 테스트 데이터 구조의 자동 생성

DTV 방송용 MPEG-2 TS를 구성하는 테이블들은 고정된 구조를 가지고 있지 않다. 하부 테이블이 삽입되거나 일정 부분이 반복적으로 확장되는 경우가 있기 때문이다. 따라서 테이블에 삽입되는 하부 테이블이나 반복 부분을 모두 하부 테이블로 구분하여 이들간의 계층 구조를 Meta data DB\_II(그림 1의 ②-3)라는 데이터 베이스로 구축하였다. Meta data DB\_II의 테이블 간의 계층 구조의 표현은 그림 2와 같다. 그리고 하부 테이블이 상부 테이블에 삽입되는 개수 또한 가변적이기에 삽입개수를 결정하는 함수를 구현 하였으며 이 함수의 고유번호를 데이터 베이스에 기록하였다.



그림 2. Meta data DB\_1의 테이블 간의 계층 구조

따라서 TDGT는 Meta data DB\_II를 참조하고 테이블 삽입 개수 결정 함수를 호출하여 테스트 데이터의 구조를 자동 생성한다.

2.3 테스트 데이터 값의 자동 생성

TDGT는 테스트 데이터의 구조를 자동 생성한 후 생성된 테이블 구조의 레코드를 하나씩 읽어서 해당 필드의 테스트 데이터 값을 생성한다. 이때 TDGT는 테스트 데이터 값을 MPEG-2 표준에 맞게 자동 생성 하기 위해 MPEG-2 표준을 데이터 베이스로 구축한 Meta data DB\_I(그림 1의 ②-2)을 참조한다. 이때 참조되는 데이터 베이스는 Meta data DB\_II와 동일한 이름의 테이블로 구성되며 그 구조는 표 1과 같다.

필드명	속성	의미
field_name	integer	MPEG-2 표준의 필드명
attribute_type	text	테스트 데이터의 생성 순서
type	integer	MPEG-2 표준의 타입
bits	integer	MPEG-2 표준의 bit 수
valid_min	integer	MPEG-2 표준의 최대 유효값
valid_max	integer	MPEG-2 표준의 최소 유효값
fixed_value	integer	MPEG-2 표준의 고정값
function_id	integer	자동 생성 함수의 고유 번호

표 1. Meta data DB\_1의 구조

표 1에서 MPEG-2 표준의 내용이 아닌 것이 attribute\_type과 function\_id이다. attribute\_type은 MPEG-2 표준에 속하는 필드의 속성을 4가지로 구분한 것으로 필드 값의 생성 순서를 가리킨다. attribute type의 의미는 표 2와 같다. function\_id는 MPEG-2 표준의 규칙이나 제약 사항 등을 기반으로 구현된 테스트 데이터 값 자동 생성 함수의 고유 번호이다. 이 함수는 Meta data DB\_II내의 type, bits, valid\_min, valid\_max, fixed\_value 등의 값을 참조하여 random selection 방식으로 테스트 데이터를 생성한다.

attribute_type	의미
type 1	자동 생성 될 수 없는 데이터. Raw data DB로 구축되어 있다.
type 2	type 1의 값에 의존적인 데이터. type 1의 값이 결정되면 그에 따라 값이 결정된다.
type 3	type 1의 값에 상관없이 MPEG-2 표준의 규칙이나 제약 사항만으로 값을 정할 수 있는 데이터.
type 4	항상 값이 고정된 데이터.

표 2. attribute\_type의 의미

3. TDGT의 분석

본 논문에서 제안한 TDGT는 DTV 방송을 위한 TS를 처리하는 DASE 시스템을 테스트 하기 위한 테스트 데이터를 생성하는 도구이다. 따라서 테스트 데이터는 MPEG-2 TS를 구성하는 테이블의 형식을 따라야 한다. MPEG-2 TS는 약 20 개의 상위 테이블과 수많은 하위 테이블로 구성되며 각 테이블은 수십 개의 필드로 구성된다. 따라서 테스트 데이터를 생성하기 위해서는 적절한 테이블을 선택하고 이 테이블 내의 필드의 값을 MPEG-2 표준에 맞게 생성해야 한다. 이러한 일련의 작업은 MPEG-2 표준에 대한 전문적인 지식이 없이는 수행하기가 매우 어려울 뿐만 아니라 테스트 데이터의 양이 매우 많기에 상당한 노력과 시간이 걸린다.

TDGT는 입력창을 통해서 테스터가 원하는 종류의 테스트 데이터에 대한 기본적인 정보를 입력 받은 후 테스터의 간섭 없이 테스트 데이터에 대한 계획을 세우고 구조와 필드 값을 MPEG-2 표준에 맞추어 수분내에 자동 생성한다. 따라서 전문적인 지식이 없는 이도 충분히 쉽게 테스트 데이터를 생성할 수 있으며 테스트 데이터를 생성하는 데 걸리는 시간과 노력을 상당히 단축할

수 있다. 또한 생성된 데이터에 대한 정보를 시나리오 파일을 통해 제공하며 커버리지 결과를 다양한 그래프를 통해 시각적으로 제공하여 테스트의 이해를 용이하게 한다.

TDGT는 다양한 종류의 테스트 데이터를 생성할 수 있다. TDGT는 테스트로부터 5가지 종류의 입력사항을 받아들이며 이를 기반으로 방송 시간, 방송 내용 등을 임의로 결정한다. 방송 내용은 앞서 설명한 Raw data DB(그림 1의 ②-1)로부터 선택하는데 Raw data DB는 테스트가 원하는 대로 업데이트를 할 수 있다. 따라서 선택할 수 있는 방송 내용의 종류는 고정되어 있지 않으며 Raw data DB의 내용이 증가할수록 TDGT가 생성할 수 있는 테스트 데이터의 종류도 증가한다. TDGT가 생성할 수 있는 테스트 데이터의 수는 그림 3과 같다.

User input 변수	TDGT 내부 변수
*service의 type ----- 24	*방송 시간의 범위 ----- 4
*syntax의 종류 ----- 2	*방송 시간 구성 ----- n
*protocol의 종류 ----- 23	*Raw data DB의 service의 종류 --- m
*DASE module의 선택 ----- 15	
*duration의 유무 ----- 2	
→ 24 * 2 * 23 * 15 * 2 = 132480	→ 4 * n * m
총 생성 개수 : 132480 * n * m 개	

그림 3. TDGT 생성 테스트 데이터의 수

MPEG-2 TS 표준은 현재 확정되지 않은 상태로 계속 수정, 확장 되고 있다. 따라서 TDGT내에 구현된 MPEG-2 표준에 관한 정보도 수정이 요구될 수 있다. TDGT는 MPEG-2 표준에 관한 정보를 Meta data DB로 분리하여 구축하였기에 수정이 용이한 구조를 가지고 있다.

4. Prototype

본 논문에서 구현한 TDGT는 MS Windows 98환경에서 JDK1.2.1(Java Development Kit)로 구현하였으며 Raw data DB, Meta data DB, Testdata DB, Analysis data DB 등의 데이터 베이스는 MS access 2000을 이용하여 구축하였다. 자바에서 데이터 베이스 접속 방법으로는 JDBC(Java DataBase Connectivity)를 사용하였다.

그림 4는 TDGT의 동작 흐름을 나타낸 것이며 그림 5는 그림 4에 맞추어 TDGT가 실제 수행되는 예를 화면으로 보인 것이다.

테스터가 자신이 원하는 테스트 데이터의 정보를 입력하면 TDGT는 테스트가 선택한 사항을 확인시켜 주며 동시에 테스트가 원하는 종류의 테스트 데이터의 예상 신뢰도를 계산하여 보여준다(그림 5의 ①,②). 이때 예상 신뢰도란 TDGT가 보유하고 있는 Raw data DB 내의 자료에 근거하여 테스트가 원하는 테스트 데이터를 생성했을 때 생성된 테스트 데이터가 얼마나 테스트의 요구사항에 부합하는지를 수치로 나타낸 것이다. 이후 테스트는 TDGT에게 테스트 데이터를 생성하도록 하거나, 예상 신뢰도가 너무 낮은 경우 Raw data DB를 업데이트 할 수 있다(그림 5의 ③,④).

TDGT는 테스트 데이터를 자동 생성한 후 테스트 데이터 스크립트 파일, 테스트 데이터 시나리오 파일, 커버리지 결과의 3가지 결과물을 테스트에게 제공한다. 테스트 데이터 스크립트 파일은 생성된 테스트 데이터를 텍스트 파일 형태로 만든 것이며 테스트 데이터 시나리오 파일은 테스트 데이터가 생성되는 과정에서 TDGT가 결정한 사항들을 기록한 텍스트 파일이다. 테스트 데이터 스크립트 파일과 시나리오 파일은 TDGT가 제공하는 에디터(그림 5의 ⑤, ⑦)로 볼 수 있으며 테스트 데이터에 대한 커버리지는 다양한 그래프를 이용하여 시각적으로 보여진다(그림 5의 ⑥).

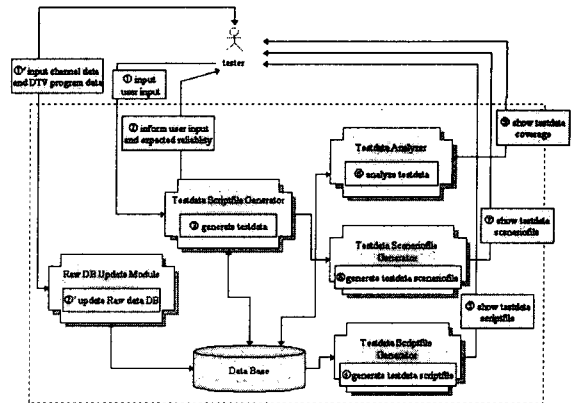


그림 4. TDGT의 동작 흐름

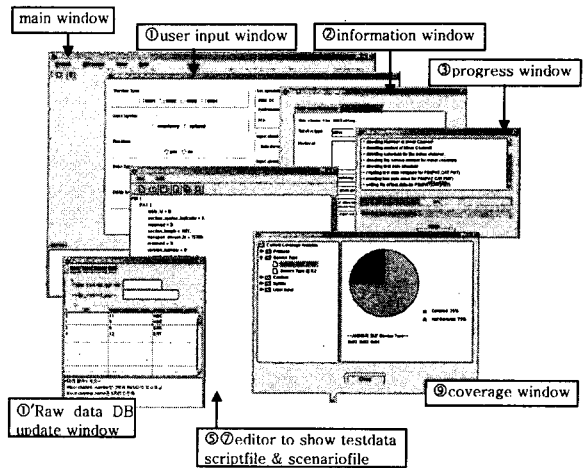


그림 5. TDGT의 수행 화면

5. 결론 및 향후 연구 과제

멀티미디어 표준인 MPEG-2 표준을 대상으로 한 테스트 데이터 자동 생성기인 TDGT를 구현하였다. 본 논문에서 제시한 TDGT는 사용자로부터 기본적인 입력 사항만을 받아 DASE 시스템을 테스트 할 수 있는 테스트 데이터를 자동 생성하는 도구이다. TDGT를 이용하면 방대하고 복잡한 테스트 데이터를 생성하는 데 드는 시간과 비용, 노력을 상당히 절감할 수 있다.

향후 fault-injection 기법을 적용시켜 다양한 오류가 삽입된 테스트 데이터를 생성할 수 있는 기능을 추가할 예정이다.

5. 참고 문헌

- [1]정재창, "최신 MPEG", 교보문고, 1995
- [2]ISO/IEC 13818-1 | ITU-T Rec. H.222.0, Information Technology — Generic coding of moving pictures and associated audio — Systems, 1996
- [3]ATSC Standard T3/S13 — ATSC Data Broadcast Standard (Draft), 2000
- [4]ATSC Standard A/65 — Program and System Information Protocol for Terrestrial Broadcast and Cable, 1997
- [5]ISO/IEC 13818-6 — MPEG-2 Digital Storage Media command & Control, Chapter 2, 4, 5, 6, 7, 9 and 11, 1998
- [6]ATSC Standard T3/S13 - ATSC Data Broadcasting Implementation Guidelines Draft Description (Doc.011), 1999