

편재형 컴퓨팅 환경에서 XHTML과 VoiceXML간 EXI 문서의 변환시스템 설계와 구현

신 경 희*, 곽 동 규 **, 유 재 우***

Design and Implementation of a XHTML to VoiceXML Converter based on EXI in Pervasive Environments

Kyoung-Hee Shin *, Donggyu Kwak **, Chae-Woo Yoo ***

요 약

편재형 컴퓨팅 환경은 다양한 컴퓨팅 장비의 접속만큼이나 그 응용분야도 다양하다. 이러한 컴퓨팅환경에서 XML은 최적의 데이터 표현이다. W3C 표준인 XML은 XSLT를 이용하여 다른 응용분야에 적절한 데이터 변환이 가능하다. 그러나 XML은 텍스트 기반으로 이진데이터 저장방식에 비해 파일의 크기가 커서 컴퓨팅 파워가 적은 편재형 컴퓨팅 환경에서는 다루기가 어렵다는 단점을 갖고 있다. 본 논문에서는 XML문서의 인코딩 방법에 대하여 살펴보고 자원을 적게 사용하는 편재형 컴퓨팅 환경에서 인코딩된 EXI 포맷의 XML문서가 다른 응용의 적합한 EXI 포맷 XML 문서로서의 변환방법을 제안한다. 다양한 응용환경 중에서 본 논문에서는 EXI 포맷의 XHTML 문서를 EXI 포맷 VoiceXML로 변환하는 시스템을 제시한다. 이 시스템은 자원이 적은 편재형 컴퓨팅 환경에서 EXI 포맷 XML 문서의 재사용성을 높이고 EXI 포맷의 XML활용에 기여할 수 있다.

Abstract

In a pervasive environment, there are various applications as much as connections of various devices. In this computing environment, XML is the most suitable data representation method. XML is able to transform data for other application areas using XSLT. XML is text-based, the file size of XML document is bigger than binary data file. Therefore, XML has a disadvantage that it is hard to deal with XML in a pervasive environment. In this paper, we survey encoding methods of XML documents, and then we propose a transform method that transforms an encoded EXI format XML document into an EXI format XML document suited for other applications. Among various applications, we present a system that transforms an EXI format XHTML document into an VoiceXML document. This system can improve reusability of EXI format XML documents in a pervasive environment and it is expected to contribute utilization of EXI format XML documents.

▶ Keyword : EXI, 이진 XML(Binary XML), XHTML, VoiceXML, 편재형(pervasive)

• 제1저자 : 신경희 교신저자 : 곽동규

• 투고일 : 2009. 08. 14, 심사일 : 2009. 10. 28, 게재확정일 : 2009. 11. 03.

* 숭실대학교전산원 멀티미디어학과 교수 ** 숭실대학교 컴퓨터학과 박사과정 *** 숭실대학교 컴퓨터학과 교수

※ 본 연구는 숭실대학교 교내 연구비 지원으로 이루어졌음.

1. 서론

무선기술과 인터넷 융합으로 접속 가능한 컴퓨팅 장비가 점차 늘어나면서 컴퓨터 환경도 다양해지고 그에 따른 응용분야도 다양해지고 있다. 이러한 편재형 컴퓨팅 환경(Pervasive Computing Environment)에서 융통성 있게 데이터를 주고 받기 위해서는 메타 데이터(Meta Data) 기반의 XML을 많이 사용하고 있다.

W3C 표준인 텍스트 기반의 XML (eXtensible Markup Language)은 데이터를 표현하는 표준으로 다양한 인터넷 응용에서 사용되고 있다. XML은 텍스트 문서에 태그를 이용하여 메타 데이터를 작성하는 방법을 사용하고 있어 어플리케이션 환경에 독립적인 장점을 갖고 있는 반면 파일의 크기가 큰 단점을 가지고 있다.

XML 파일의 크기가 큰 단점을 해결하기 위하여 XML 형태는 유지하면서 파일의 크기를 줄이는 여러 가지 방법이 연구되었다(1)(2)(3)(4)(5)(6). 그 중 EXI (Efficient XML Interchange)(6) 인코딩방식은 W3C (World Wide Web Consortium)의 표준으로 라이선스 비용이 없고 SAX (Simple API for XML)(7) 파싱 인터페이스를 이용하여 많은 응용에서 사용할 수 있다. 휴대용 단말기와 같이 자원이 제한되었거나 적은 대역폭이 요구되는 환경인 경우 XML 문서를 전송할 때는 인코딩된 XML EXI 포맷을 이용하는 것이 보다 효율적이다.

XML 데이터를 다양한 응용분야에 적합한 데이터 방식으로 표현하기 위해서는 XSLT (Extensible Stylesheet Language Transformations)(8) 처리가 필요하고, 다양한 XML 응용분야에서 빠른 데이터 전달이 요구되는 경우는 EXI 포맷 형태가 필요하다. 따라서 편재형 컴퓨팅 환경에서 EXI 포맷 정보를 각 응용분야에 적합한 단말기에 전달되기 위해서는 먼저, EXI 포맷의 XML 정보를 텍스트 기반의 XML 문서로 변환하고 이를 XSLT를 이용하여 목표 단말환경에 적합한 XML 문서로 변환 생성한 후, 이를 다시 EXI 포맷의 XML로 변환되어야 한다.

경량의 휴대용 접속장치와 같은 단말기가 사용되는 컴퓨팅 환경에서 인코딩되지 않은 텍스트 기반의 XML 문서는 파일 크기 때문에 프로세싱이 어렵다. 따라서 본 논문에서 다양한 단말기가 사용되는 편재형 컴퓨팅 환경에서 좀 더 빠른 데이터 전달을 위하여 텍스트 기반의 XML 문서를 EXI 포맷을 이용하여 다양한 응용환경에서 사용될 수 있는 변환방법을 제안한다. EXI 포맷으로 변환된 XML 문서는 텍스트 XML 문

서를 위한 저장 공간이 필요하지 않고 실행 시간도 단축시킬 수 있다.

본 논문에서는 EXI 포맷으로 저장되어 있는 XHTML (eXtensible Hyper Text Markup Language)(9) 문서를 EXI 포맷의 VoiceXML(10) 문서로 변환하는 시스템을 제안한다. XHTML은 웹 브라우저(Web Browser)를 통해 사용자에게 정보를 제공하고 VoiceXML은 전화기와 같은 음성 브라우저를 통해 사용자에게 정보를 제공하는 각기 다른 XML 응용이며 W3C의 표준이다. XHTML형태로 제공되는 정보는 웹 브라우저를 이용하여 정보를 수취해야 하는데 사용자 단말기에 웹 브라우저가 없을 경우 정보 취득이 곤란하다. 그 대신 음성 브라우저가 지원되는 경우라면 XHTML 문서를 VoiceXML 문서로 변환시키는 작업이 이루어진 후에 그 음성 브라우저를 이용하여 정보를 취득하게 된다. 이러한 문서변환 작업을 특히 EXI 포맷으로 작성된 문서를 텍스트 기반의 XML 문서로 변환하는 것과 텍스트기반의 XML문서를 EXI포맷으로 변환하는 작업은 단말기 환경에서 실행된다면 각 단말기들의 제한적인 컴퓨팅 파워 때문에 부하가 많이 걸리게 된다.

본 논문에서 제안하는 EXI 변환시스템은 EXI 파서, 변환 규칙 적용기, EXI 생성기인 세 개의 모듈로 구성된다. EXI 파서는 바이너리로 인코딩되어 있는 XML 문서를 SAX와 같은 방법으로 파싱하는 모듈이고, 변환 규칙 적용기는 XML응용분야에 따른 변환 규칙을 적용하는 모듈이고, EXI 생성기는 XML 문서를 EXI 포맷으로 생성하는 모듈이다. 이러한 모듈 구성은 EXI 포맷의 XML문서가 다양하게 응용되는 편재형 컴퓨팅 환경에 적용하기 용이하고 적은 컴퓨팅 파워를 요구하는 EXI 포맷의 XML 활용에 기여할 수 있고 데이터 재사용성을 증가시킬 수 있다는 장점을 갖는다.

논문 구성은 2장에서 관련연구와 기존의 XSLT를 이용한 XML 변환 방법에 관해 논한 후 3장에서 제안하는 변환시스템 구조를 설명한다. 4장에서는 기존의 XSLT를 이용한 변환과 본 논문에서 제안하는 시스템에서의 변환을 비교하여 실험한 후 5장에서 결론과 향후 연구 과제에 대하여 기술한다.

II. 관련 연구

2.1 이진 XML과 EXI

XML은 텍스트를 기반으로 하는 데이터 표현 방법으로 컴퓨팅 환경과 독립적으로 사용할 수 있어 다양한 응용에서 사용하고 있다. 하지만 텍스트를 기반으로 하고 있어 이진 데이터를 사용하는 경우보다 파일이 큰 단점을 가진다. 이러한 단

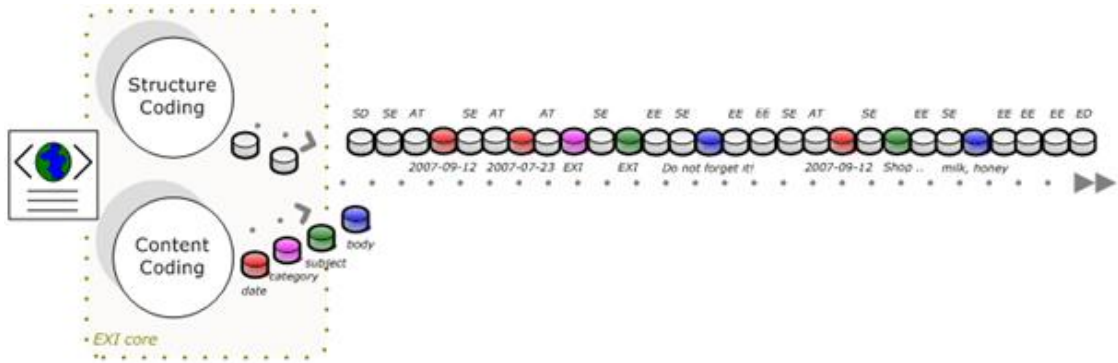


그림 1. EXI 스트리밍
Fig. 1. EXI Streaming

점을 극복하고자 텍스트를 기반으로 하는 XML 문서를 이진화하여 크기를 줄이고자 하는 연구가 진행되었다.

XML을 이진화하는 방법은 크게 세 가지로 구분한다. 첫 번째 방법은 XML 문서 전체를 압축하는 방법이다[1]. 이 압축방법은 문서를 파싱하기 위해 압축을 풀어 기존의 XML 문서로 복구하여 사용하는 방법으로 XML 전송에는 유리하지만 파싱을 위한 비용이 많이 들어 잘 사용하지 않는다. 두 번째 방법은 문서의 구조를 분리하여 콘텐츠(Content)와 인덱스(Index)를 따로 관리하고 이중콘텐츠를 압축하는 방법이다[2][3]. 이 방법은 문서 전체를 압축하는 방법에 비해 문서 구조를 파악할 수 있는 장점을 갖지만 XML 문서가 전부 필요할 경우 압축 방식과 유사한 비용이 든다는 단점을 갖는다. 세 번째 방법은 본래의 XML 문서의 의미와 구조를 유지하여 고유의 인코딩 포맷으로 정의하는 방법으로 XML 문서의 데이터들을 각각의 데이터 타입에 합당한 낮은 수준의 형식을 사용한다[4][5][6]. 이 방법은 XML 구조를 유지하며 데이터를 저장하기 때문에 기존의 XML 파싱 방법을 유사하게 사용할 수 있는 장점을 가진다. 세 번째로 인코딩을 사용하는 이진 XML 중 EXI[6]는 W3C의 표준으로 라이선스에 대한 부담이 없다. 그리고 텍스트 기반의 XML 포맷과 문서 구조가 유사하여 SAX나 DOM(Document Object Model)과 같이 기존의 XML 파싱 방법을 유사하게 사용할 수 있는 장점을 가진다. 그림 1은 EXI 포맷의 스트리밍(Streaming) 구조를 보인다.

그림 1에서 상단에 이름(SD, SE, AT, EE, ED)이 있는 블록(block)은 XML 문서의 구조를 인코딩한 블록을 나타내고 그 나머지부분은 콘텐츠를 인코딩한 블록을 표현하고 있다. XML의 문서 구조를 인코딩한 블록은 SD (Start Document)와 SE (Start Element), AT (Attribute), EE (End Element), ED (End Document)로 구성되어 있는데 이들이 등장하는 순서는 텍스트 기반의 XML 문서 구조와 유사하다. 그러므로 기존의 XML 파싱 방법을 유사하게 사용할 수 있다.

2.2 XSLT를 이용한 EXI 포맷 XML 문서의 변환

XSLT는 XML의 한 응용으로 XML 문서의 스타일을 정의하는 언어이다[8]. XSLT는 XML 문서를 다른 형태의 문서로 변환할 변환 규칙을 기술한다. 따라서 어떤 응용분야에서 사용되고 있는 XML 문서를 다른 응용분야에서 사용하기 위해서는 해당 분야에 적합하도록 XML 문서를 변환해야 한다. 또한, 텍스트 기반의 XML의 경우 변환 규칙과 시스템에 관한 연구가 진행되었다[11][12][13]. 하지만, XSLT는 텍스트 기반의 XML을 지원하고 있어 EXI 포맷에서 사용하기 위해서는 EXI 포맷으로 작성된 XML 문서를 텍스트로 변환한 후 다시 EXI 포맷으로 변환해야 한다.

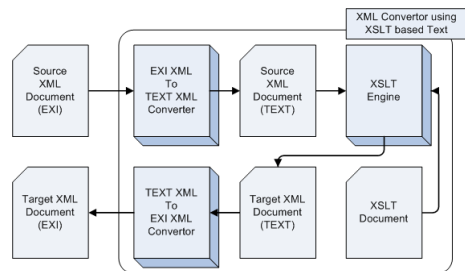


그림 2. 텍스트 기반의 XSLT를 이용한 XML 변환
Fig. 2. XML Converting using XSLT based Text

그림 2는 EXI 포맷정보를 XSLT를 이용한 목표 EXI 포맷정보로의 변환을 보인다. 그림 2와 같이 XSLT를 이용하여 EXI 기반의 XML을 변환할 경우 EXI XML문서를 텍스트 XML 문서로 변환하는 변환기(EXI To TEXT XML Converter), 목표 응용분야에 적합한 XML문서로 변환시키는 XSLT처리기(XSLT Engine), 텍스트 XML문서를 EXI 포맷으로 변환하는 변환기(TEXT to EXI XML Converter)가 필요하다. 이는 편제형 환경에서 자원이 적은 단말기를 사용할 경우 처리 속도의 저하를 가져올 수 있다.

III. EXI 포맷 기반의 XHTML과 VoiceXML 문서 변환 시스템

XHTML은 웹 브라우저를 통해 사용자에게 정보를 제공하기 위한 마크 업 언어로서 사용자에게 화면을 통해 정보를 제공한다. 편재형 컴퓨팅 환경에서 사용자는 PDA와 같은 웹 브라우저를 사용할 수 있는 휴대용 단말기를 통해 XHTML에 작성되어 있는 정보를 수취할 수 있다. 이 때 XHTML 문서를 EXI 포맷으로 제공하면 적은 대역폭과 컴퓨팅 자원으로 XHTML 문서의 정보를 취득할 수 있다. 하지만 사용자 단말기의 화면이 작거나 프로그램을 지원하지 않아 웹 브라우저를 사용할 수 없으면 사용할 수 있는 응용에 합당한 형태의 XML로 변환해야 한다. 그림 3은 XHTML로 제공되는 정보를 전화기와 같은 단말기에서 사용하기 위한 환경을 보인다.

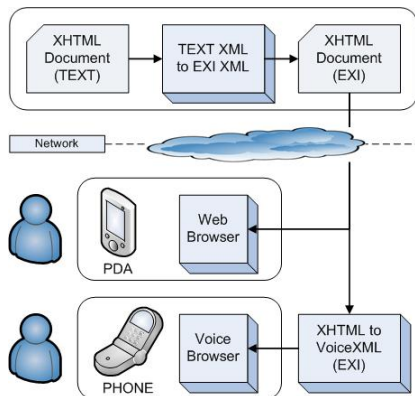


그림 3. EXI 기반 XHTML을 VoiceXML
Fig. 3. XHTML to VoiceXML based EXI

본 논문에서는 EXI 포맷의 XHTML 문서 정보를 전화기와 같은 웹 브라우저를 사용하기 어려운 단말기에서 수취하기 위해 EXI 포맷의 VoiceXML로 변환하는 시스템에 대하여 설명한다. 일반적으로 XML의 변환은 XSLT를 이용한다. 하지만 EXI 포맷으로 작성된 XML 문서를 다른 응용에서 사용하는 EXI 포맷의 XML 문서로 변환하기 위해서는 EXI 포맷으로 되어 있는 XML 문서를 텍스트 기반의 XML 문서로 변환한 후 XSLT를 적용해야 한다.

본 논문에서 제안하는 시스템은 서버에서 제공해 주는 EXI 포맷의 XHTML 문서를 텍스트 문서로 변환하지 않고 EXI 포맷의 VoiceXML 문서로 직접 변환한다. 그림 4는

EXI 기반의 XHTML을 VoiceXML로 변환하는 시스템의 구조를 보인다. 이 시스템은 프로그램을 경량화할 수 있고 변환에 요구되는 시간을 줄일 수 있는 장점을 가진다. 변환 프로그램의 경량화는 자원이 적은 단말기에 적용하기 용이하다.

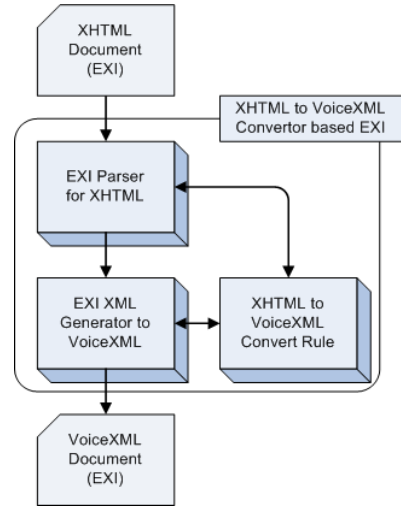


그림 4. EXI 포맷 기반 XHTML을 VoiceXML 문서 변환시스템 구조
Fig. 4. A Converting System of XHTML to VoiceXML based EXI format

텍스트 기반의 XML문서를 생성하지 않고 EXI포맷간에 XHTML 문서를 VoiceXML 문서로 변환하는 제안시스템은 EXI 포맷 XHTML 파서(EXI Parser for XHTML), EXI 포맷의 VoiceXML 문서 생성기(EXI XML Generator to VoiceXML), VoiceXML으로의 변환규칙 적용기(XHTML to VoiceXML Convert Rules)로 구성되어 있다(그림 4).

EXI 포맷 XHTML 파서는 EXI ficent[14]를 이용한다. EXI ficent는 sourceforge.net[15]을 통해 진행되고 있는 오픈소스 프로젝트(Open Source project)로 EXI 포맷으로 작성된 XML 문서를 DOM[16]이나 SAX[7]를 이용하여 파싱할 수 있고 EXI 포맷으로 문서를 인코딩 할 수 있는 방법을 제공한다. EXI 포맷 XHTML 파서는 전체 시스템의 입력인 EXI 포맷의 XHTML문서를 SAX 이벤트를 생성한다. EXI 포맷 VoiceXML 문서 생성기는 XHTML을 VoiceXML로 변환하는 규칙에 따라 문서를 생성하는 것으로 EXI ficent를 이용한다. VoiceXML으로의 변환규칙 적용기는 XHTML의 태그(tag)와 VoiceXML의 태그를 매핑(Mapping)하는 규칙을 가지고 있다.

이벤트 기반의 파싱 방법으로 처리되는 본 논문의 시스템은 문서의 변환 규칙은 해당 이벤트에 따라 EXI ficent를 이용하여 인코딩한다. 변환시스템은 표 1의 알고리즘으로 처리된다.

표 1. XHTML to VoiceXML 변환 시스템의 알고리즘
Table 1. Algorithm of Converting System for XHTML to VoiceXML

```

SD : Start Document
SE : Start Element
AT : Attribute
EE : End Element
ED : End Document

입력 : xhtmlDecoder // EXI 포맷 XHTML
출력 : voiceXMLEncoder // EXI 포맷 VoiceXML

converterXHTMLtoVoiceXML(xhtmlDecoder){
voiceXMLEncoder SD 인코딩;
for(int i = 0; i < xhtml의 블록 개수; i++){
EXIBlock block = xhtmlDecoder.getBlock(i) // XHTML 문서
의 i번째 EXI 블록
switch(block){ // 매핑 규칙에 따라 인코딩
case xhtmlStartElement :
voiceXMLEncoder vxml SE 인코딩;
voiceXMLEncoder vxml SE 어트리뷰트 인코딩;
case headStartElement :
voiceXMLEncoder form SE 인코딩;
voiceXMLEncoder form AT 인코딩;
voiceXMLEncoder block SE 인코딩;
voiceXMLEncoder prompt SE 인코딩;
case bodyStartElement :
...
case xhtml End Element :
voiceXMLEncoder vxml EE 인코딩;
case head End Element :
voiceXMLEncoder form EE 인코딩;
voiceXMLEncoder block EE 인코딩;
voiceXMLEncoder prompt EE 인코딩;
case body End Element :
...
}
voiceXMLEncoder ED 인코딩;
}
    
```

XHTML은 화면과 웹 브라우저(Web Browser)를 통해 시각적으로 사용자에게 정보를 제공한다. 반면에 VoiceXML은 전화기와 음성 브라우저를 통해 청각적으로 사용자에게 정보를 제공하기 위한 XML의 한 응용이다. 두 응용의 정보 전달 매체를 비교하면 XHTML은 시각적인 매체로 시각적인 제약 없이 받는 반면 VoiceXML은 청각적인 매체로 시간적인 제약을 받는다. 그러므로 XHTML 문서의 크기가 VoiceXML 문서의 크기에 비해 상대적으로 많은 양의 정보를 가질 수 있다. 이와 같은 문서의 정보량 차이로 인해 서로 다른 특성을 갖는 응용간의 문서 변환은 문서를 분할하거나 병합하여 변환해야 한다[11][12][13][17]. 본 논문의 시스템도 필요에 따라 기존 XHTML문서의 정보가 많을 경우 정

보를 분할할 수 있다. XML 문서를 분할하는 방법은 태그의 패턴을 분석하는 방법[17] 그리고 통계적인 방법[11], XML 스키마(Schema)를 분석하는 방법[12] 등이 있다.

IV. 실험 및 성능 평가

실험방법은 본 논문에서 제안한 EXI 포맷으로 작성된 XHTML 문서를 EXI 포맷의 VoiceXML로 변환 방법과 일반적으로 텍스트 문서로 변환한 후 XSLT를 이용하여 XHTML문서를 VoiceXML문서로의 변환 방법을 구분하여 각 성능을 비교 평가하였다. 표 2는 실험 환경이다.

표 2. 실험 환경
Table 2. Environment of Testing

항목	버전
프로세서	Intel(R) Pentium(R) M processor 1.86GHz
메모리	1.86GHz 2GB RAM
운영체제	Windows XP Professional Version 2002 Service Pack 3
개발환경	Java™ 2 Runtime Environment, Standard Edition 1.6.0_13

표 2와 같은 실험 환경에서 각 변환방법에 따른 성능을 평가하기 위해 표 3과 같은 예제의 XHTML 문서를 EXI 포맷으로 작성하고 EXI 포맷의 VoiceXML 문서 표 4를 생성한다.

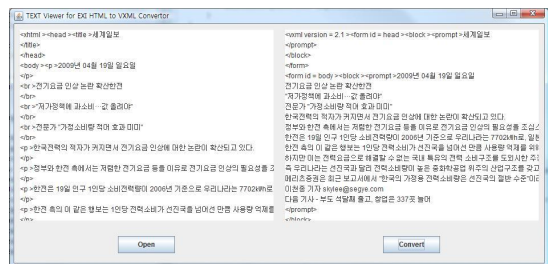


그림 5. XHTML to VoiceXML for EXI 프로그램
Fig. 5. XHTML to VoiceXML Program for EXI

그림 5는 텍스트 환경의 프로그램 실행 화면이다. 그림에서 왼쪽은 표 4에 해당하는 문서이고 오른쪽은 표 5에 해당하는 문서이다.

표 3. 변환 대상 XHTML 문서의 텍스트 포맷
Table 3. Source XHTML Document based Text Format

```
<?xml version="1.0" encoding="euc-kr"?>
<xhtml><head>
<title>세계일보</title>
</head><body>
<p>2009년 04월 19일 일요일</p>
<br>전기요금 인상 논란 확산한전</br>
<br>"저가정책에 과소비...값 올려야"</br>
<br>전문가 "가정소비량 적어 효과 미미"</br>
<p>한국전력의 적자가 커지면서 전기요금 인상에 대한 논란이 확산되고 있다.</p>
...
<p>이천중 기자 skylee@segye.com</p>
<a href="news2.html">다음 기사 - 부모 석달째 줄고, 창업은 337곳 늘어</a>
</body>
</xhtml>
```

각 문서들의 변환 규칙은 그 특성에 따라 작성되어야 한다. 예를 들어 XHTML 문서를 사용하는 웹 브라우저는 마우스를 이용한 링크로 페이지를 이동하고 VoiceXML 문서를 사용하는 전화기와 음성 브라우저는 전화기의 숫자 버튼을 이용하여 다른 페이지의 정보로 이동한다.

표 4. 변환 후 VoiceXML 문서의 텍스트 포맷
Table 4. Target VoiceXML based Text Format

```
<?xml version="1.0" encoding="EUC-KR"?>
<block><prompt>세계일보</prompt></block></form>
<form id="body"><block><prompt>2009년 04월 19일
<?xml version="1.0" encoding="EUC-KR"?>
<vxml version="2.1"><form id="head">
<block><prompt>세계일보</prompt></block></form>
<form id="body"><block><prompt>2009년 04월 19일 일요일
전기요금 인상 논란 확산한전"저가정책에 과소비...값 올려야"전문가
"가정소비량 적어 효과 미미"한국전력의 적자가 커지면서 전기요금
인상에 대한 논란이 확산되고 있다.
...
이천중 기자 skylee@segye.com<form id="link">
<field id="next" type="digits?length=1"><prompt>다음 기
사 - 부모 석달째 줄고, 창업은 337곳 늘어난 1번을 누르세
요.</prompt>
<grammar xml:lang="en-US" root="TOP" mode="dtmf">
<rule id="TOP" scope="public"><one-of><item> 1 <tag>
&lt;next "nextArticle"&gt; </tag></item>
</one-of></rule></grammar>
<filled><if cond="next == 'nextArticle'">
<goto next="news2.vxml"/></if></filled>
</field></form></prompt></block></form></vxml>
```

본 논문에서 제안한 변환 시스템과 텍스트 기반의 XSLT 시스템의 성능을 비교 평가하기 위해서 표 5와 같은 XSLT 문서를 사용한다. 이 문서는 XHTML 포맷의 신문기사의 내용을 VoiceXML 포맷으로 변환하여 음성 인터프리터를 이용하는 사용자에게 제공할 수 있도록 작성된 문서이다.

표 5. XHTML to VoiceXML XSL 문서
Table 5. XSL Document of XHTML to VoiceXML

```
<?xml version="1.0" encoding="euc-kr" ?>
<xsl:stylesheet version="1.0"
xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:template match="xhtml">
<vxml version="2.1"><xsl:apply-templates/></vxml>
</xsl:template>
<xsl:template match="head"><form id="head"><block>
<prompt><xsl:apply-templates/></prompt>
</block></form></xsl:template>
<xsl:template match="body"><form id="body"><block>
<prompt><xsl:apply-templates/></prompt>
</block></form></xsl:template>
<xsl:template match="a"><form id="link">
<field id="next" type="digits?length=1">
<prompt><xsl:value-of select="."/>는 1번을 누르세요.</prompt>
<grammar xml:lang="en-US" root="TOP" mode="dtmf">
<rule id="TOP" scope="public"><one-of>
<item> 1 <tag> <![CDATA[ <next "nextArticle" ]> ]>
</tag></item></one-of></rule>
</grammar><filled><if cond="next == 'nextArticle'">
<goto next="news2.vxml"/></if></filled>
</field></form></xsl:template>
<xsl:template match="br">
<xsl:value-of select="."/></xsl:template>
<xsl:template match="p">
<xsl:value-of select="."/></xsl:template></xsl:stylesheet>
```

표 6은 텍스트기반의 XML정보를 XSLT로 변환시켜 인코딩하는 방법[10][17]과 본 논문에서 제안한 변환시스템에 따른 방법에 대한 실험한 결과이다.

표 6. 성능 평가 결과
Table 6. Performance Evaluation

구분	텍스트 기반의 XSLT를 이용한 변환	제안하는 시스템	개선율(%)
객체사용량 (object)	14,800	8,512	57.51
메모리사용량 (KB)	1,126.4	430.2	38.19
실행 시간 (ms)	542.3	297	54.77

실험에 따른 변환 시스템의 성능 평가는 객체 사용량과 메모리 사용량, 실행 시간으로 구분하였고 평가를 위해 이클립스(Eclipse)[18]의 플러그인으로 제공되는 MAT(Memory Analyzer Tool)[19]을 이용한다. 표 6의 평가 결과에서와 객체사용량은 57%로 감소하였고 메모리 사용량도 38%로 줄었으며 실행시간이 54%로 빨라졌다. 따라서 본 논문에서 제안한 변환시스템은 텍스트 기반의 XSLT 이용한 변환보다 약 2배의 성능 향상을 가진다.

V. 결론 및 향후연구과제

EXI[6]는 W3C의 표준으로 XML의 형태는 그대로 유지하면서 파일을 인코딩하여 크기를 줄이는 방법이다. 이는 자원이 적은 편재형 컴퓨팅 환경에서 유용하게 사용될 수 있다. 한 응용에서 사용되는 EXI 포맷의 XML 문서를 다른 응용에서 재사용하기 위해서는 문서의 변환이 필요하다. 기존의 텍스트 기반의 XML 문서는 XSLT를 이용하여 변환할 수 있다. 하지만 인코딩된 EXI 환경에 XSLT를 적용하기 위해서는 EXI 포맷으로 작성된 문서를 텍스트로 변환해야 하고 이를 XSLT에 적용한 후 다시 EXI 포맷으로 변환해야 한다. 이는 자원에 제약사항이 있는 편재형 컴퓨팅 환경에서는 적합하지 않다.

본 논문은 편재형 컴퓨팅 환경에서 다양한 응용에 적용하기 위한 EXI 변환 방법을 제안하고 그다양한 응용 환경 중 EXI 포맷의 XHTML 문서를 EXI 포맷의 VoiceXML 문서로 변환하는 시스템을 제안하였다. 본 논문에서 제안한 시스템은 실험 평가해 본 결과 메모리 사용량과 실행 시간을 살펴 보았을 때 기존의 텍스트 기반 XSLT를 이용한 변환에 비해 약 2배 정도의 성능 향상을 보인다. 또한 본 시스템은 단일 응용에서 사용하고 있는 EXI 포맷 XML 문서를 다른 응용에 적용하기가 용이하여 데이터의 재사용성을 증가시킬 수 있다. 하지만 이 방법은 두 응용간의 변환으로 텍스트 기반의 XSLT와 같이 범용적으로 쉽게 사용할 수 있는 방법에 대한 연구가 필요하다.

참고문헌

- [1] XMill, <http://sourceforge.net/projects/xmill>.
- [2] Girardot, M., Sundaresan, N., "Efficient representation and streaming of XML content over the Internet medium," Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on Vol.1, No.30, pp.67-70, July 2 Aug. 2000.
- [3] Hongzhi Wang., Jianzhong Li., Jizhou Luo; Zhenying He., "XCpaqs: compression of XML document with XPath query support," Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference on Vol.1, No.5-7, pp.354-358, Apr. 2004.
- [4] Fast Infoset, <http://asn1.elibel.tm.fr/en/xml/#fast-infoset>.
- [5] XML Binary Information Set, <http://sourceforge.net/projects/xml>.
- [6] EXI, <http://www.w3.org/XML/EXI>.
- [7] SAX, <http://www.saxproject.org>.
- [8] XSLT, <http://www.w3.org/TR/xslt>.
- [9] XHTML, <http://www.w3.org/TR/xhtml1>.
- [10] VoiceXML, http://www.w3.org/TR/voice_xml20.
- [11] 최훈일, 장영건, "HTMLtoVoiceXML 변환기의 설계 및 구현", 한국정보과학회논문지 : 컴퓨팅의 실제, 제 7권, 제 6호, 448-569쪽, 2001년 12월.
- [12] 광동규, 최종명, 유재우, "마크업 문서의 대수적 분석에 의한 분할과 변환", 프로그래밍언어학회 논문지 제 17권, 제 3호, 51-59쪽, 2003년 11월.
- [13] Narayan Annamalai, Gopal Gupta, B. Prabhakaran, "Accessing Documents via Audio: An Extensible Transcoder for HTML to VoiceXML Conversion" ICCHP 2004, pp. 339-346, 2004.
- [14] EXI Ficient, <http://exificent.sourceforge.net>.
- [15] SOURCEFORGE, <http://www.sourceforge.net>.
- [16] DOM, <http://www.w3.org/DOM>.
- [17] David W. Embley, Y. S. Jiang, Yiu-Kai Ng, "Record-Boundary Discovery in Web Documents", SIGMOD Conference, pp. 467-478, 1999.
- [18] Eclipse, <http://www.eclipse.org>.
- [19] MAT, <http://www.eclipse.org/mat>.
- [20] 조정길, 금영옥, "XML 데이터의 제약조건 보존을 위한 변환 기법", 한국컴퓨터정보학회 논문지 제 14권, 제 5호, 1-9쪽, 2009년 5월.
- [21] 김천식, 홍유식, "텍스트 마이닝을 이용한 XML 문서 분류 기술", 한국컴퓨터정보학회 논문지 제 11권, 제 2호, 15-23쪽, 2006년 5월.

저 자 소 개



신 경 희

1988: 인하대학교
전산학과 이학사.
1996: 숭실대학교
컴퓨터학과 공학석사.
2003: 숭실대학교
컴퓨터학과 공학박사
현 재: 숭실대학교전산원
멀티미디어학과 교수.
관심분야: 프로그래밍언어, 컴파일러,
XML, 멀티미디어, 게임프
로그래밍



곽 동 규

2002: 서경대학교
응용수학과 이학사.
2004: 숭실대학교
컴퓨터학과 공학석사.
현 재: 숭실대학교
컴퓨터학과 박사과정.
관심분야: 프로그래밍언어, 컴파일러,
XML, 임베디드시스템, 유
비쿼터스



유 재 우

1976: 숭실대학교
전자계산학과 이학사.
1985: 한국과학기술원
전산학과 공학박사.
1986~87, 96~97: 코넬대학교
객원교수.
1999~2001: 한국정보과학회
프로그래밍 연구회
위원장.
현 재: 숭실대학교
컴퓨터학과 교수.
관심분야: 프로그래밍언어, 컴파일러,
XML, 임베디드시스템, 유
비쿼터스