

무선전용 다중 언어의 번역을 지원하는 변환기의 구현

최 지원[†]·김기천^{**}

요 약

유선 환경에서 시작한 인터넷에 현재 존재하고 있는 웹 콘텐츠들을 모바일 환경에서 이용할 수 있게 하기 위하여 UP.Link에서는 자동으로 HTML을 HDML로 번역해주는 기능을 가지고 있다. UP 브라우저에서 HTML 페이지를 액세스할 때 UP.Link 내의 에이전트는 번역기에 HTML과 HDML간의 번역을 처리하게끔 한다. I-Mode는 일본의 NTT-DoComo에서 개발 되었으며 이미 여러 가지 서비스를 통해 엄청난 콘텐츠를 보유하고 있다. Stinger 프로젝트에 의해 개발된 ME(Micro Explorer) 또한 많은 부가기능을 가지고 있다. 본 논문에서는 WAP에서 C-HTML과 mHTML의 콘텐츠를 수용할 수 있는 변환기를 설계하고 구현하였다. I-Mode의 C-HTML과 MS사의 ME는 HTML의 간이형으로 콘텐츠 제공 업체들이 다른 언어를 익힐 필요가 없으므로 WAP에 비해 콘텐츠 개발에 유리하다. C-HTML과 WML은 모두 모바일 환경에서 쓰이기 때문에 한 페이지 내에 제한되어 있는 용량 또한 비슷하다. 그러므로 C-HTML을 WML로, mHTML을 WML로 바꾸는 변환기를 개발하기 위해서 먼저 각 언어와 관련된 스펙을 분석하고 각각의 서로 대응되는 element를 찾아 표를 작성하고 난 후에 변환을 위한 알고리즘을 적용하여 변환기를 구현한다. 만약 대응되는 element가 없을 경우에는 WML에서 지원하는 여러 가지의 태그들을 조합하여 최대한 같은 형식으로 디스플레이 해줄 수 있는 방안을 제시한다. 편집기에서는 변환된 WML 소스를 콘텐츠 저작자가 편집 기능을 이용하여 다른 내용을 추가하거나 내용을 바꿀 수 있도록 지원한다. 구현 결과 약 90% 이상의 변환이 가능하다.

Development of Converter supporting Multi-languages for Mobile Network

Jiwon Choi[†] · Keecheon Kim^{**}

ABSTRACT

UP Link is One of the commercial product which converts HTML to HDML convertor in order to show the internet www contents in the mobile environments. When UP browser accesses HTML pages, the agent in the UP Link controls the convertor to change the HTML to the HDML. I-Mode, which is developed by NTT-DoComo of Japan, has many contents through the long and stable commercial service. Micro Explorer, which is developed by Stinger project, also has many additional function. In this paper, we designed and implemented WAP convertor which can accept C-HTML contents and mHTML contents. C-HTML format by I-Mode is a subset of HTML format. mHTML format by ME is similar to C-HTML. So the content provides can easily develop C-HTML contents compared with WAP and the other case. Since C-HTML, mHTML and WML are used under the mobile environment, the limited transmission capacity of one page is also similar. In order to develop converter for mobile device, we first surveyed on the related specifications, then we find matched element on each element to make a match table. After that, we apply conversion algorithm on it. If we can not find matched element, we arrange some tags which only can be supported by WML to display in the best shape. By the result, we can convert over 90% contents.

키워드 : 이동인터넷(Mobile Internet), WAP, I-Mode, ME(Mobile Explorer), 변환기(Converter)

1. 서 론

1990년대 들어 인터넷은 여러 분야에 획기적인 답안을 제시해 주고 있는 큰 변화라 할 수 있다. 정보, 통신, 방송 및 사회 전 분야에서 인터넷에 적용하기 위한 기술들이 속속들이 나오고 있고 유선으로부터 시작한 그 발전 방향은 생활에 더욱 큰 편리함을 제공하기 위하여 무선에서도 인터넷을 서비스하기 위한 여러 가지 방안들이 계속해서 제시되고 있다. 무선에서의 인터넷 제공을 위한 움직임들을 보면, 첫 번째로 WAP(Wireless Application Protocol)을 가장 먼저 언급하지 않을 수 없다. WAP은 현재 서비스 가능한 모바일 인터넷에서 가장 강력한 단말기 플랫폼이라 할 수 있다. 단말기는 그 특성상 제한적인 성능(CPU, Memory, Display 및 전송 속도)을 가질

수밖에 없으며, 이러한 특성에 맞게 가장 적절히 그리고 잘 정의된(well defined) 형태로서, 각종 데이터서비스 및 인터넷 접속을 지원하도록 WAP은 설계되어 있다. 특히 WML(Wireless Markup Language)은 함축된 Tag기능과 바이너리 코딩의 사용으로 낮은 전송속도와 메모리, 한정된 디스플레이기능을 갖는 현재의 단말기에 매우 적합하다. 특히 상위 Application 들이 프로그래밍하기 쉽고 통일된 형태로 동작할 수 있게끔 하는 WAE(Wireless Application Environment)이 잘 정의되어 있다. 현재 WAP에 대한 표준은 WAP Forum에서 진행 중에 있으며, 현재의 2G 뿐 아니라 앞으로 IMT-2000과 같은 3G 이동 통신에서도 사용될 수 있도록 표준이 진행 중에 있다.

WAP에 대응할 만한 단말기 플랫폼으로 I Mode를 들 수 있다. 일본 최대의 휴대폰 서비스 제공 업체인 NTT는 9.6Kbps 패킷 통신망을 사용하여 이동 전화만으로 인터넷에 접속해서 음성통화는 물론, 모바일 banking, 전자 메일, 전화번호부나 레스토랑 가이드와 같은 데이터베이스 검색, 증권 정보, 게임 뉴스

* 이 논문은 2000학년도 건국대학교 학술진흥연구비 지원에 의한 논문임.

[†] 준 회원 : 건국대학교 대학원 컴퓨터공학과

^{**} 종신회원 : 건국대학교 컴퓨터공학과 교수

논문접수 : 2002년 1월 2일, 심사완료 : 2002년 1월 16일

제공, 기차표 항공권 예약을 포함해 70여 가지 서비스를 제공하는 I-Mode를 선보였다. 현재 엄청난 가입자를 가지고 있으며 콘텐츠 개발이 용이하기 때문에 I-Mode의 가입자는 계속해서 증가하고 있고 하루에 20~30개 정도의 I-Mode 사이트가 생기는 등 빠른 성장을 보이고 있다. I-Mode는 PCD-P라 불리는 도쿄모의 패킷 망에 의해 제공되어지며 따라서 시간이 아닌 데이터의 양에 따라 금액이 정해진다. I-Mode는 HTML 3.0의 간이형인 C-HTML과 약간의 추가된 태그를 사용하여 콘텐츠 제공 업체들이 또 다른 언어를 익힐 필요가 없도록 하였다.

WAP과 W3C에 대응하여 또 하나의 무선 단말기 플랫폼으로 마이크로소프트는 CDMA 관련 원천기술을 보유하고 있는 퀵컴사와의 제휴를 통해 Wireless knowledge라는 합작회사를 설립하여 휴대폰에서 인터넷을 검색할 수 있도록 윈도CE를 근간으로 한 이동통신용 웹브라우저를 개발하여, ME(Mobile Explorer)를 서비스하고 있으며 Stinger 프로젝트를 통해 지속적인 업그레이드 과정에 있다.

최근 국내 이동 통신 업계의 움직임을 보면 WAP과 ME를 이용하여 이동 단말기에 더욱 많은 콘텐츠를 제공하기 위한 노력이 가중되고 있다. 앞서 열거한 세 가지 서비스를 비교한다면 HTML의 간소화된 형태인 C-HTML을 사용하는 I-Mode나 mHTML을 이용하는 ME가 콘텐츠 제공 업체들의 관점에서 볼 때에는 이점이 더 많다. 본 논문에서는 세 가지 서비스에서 사용되어지는 각각의 언어들의 호환을 위한 콘텐츠 변환기의 설계 및 구현을 목표로 하고 있다.[8,9]

2. 관련 기술

2.1 WAP(Wireless Application Protocol)

WAP은 현재 빠르게 진보하는 2가지 기술인 무선 데이터와 인터넷을 합쳐 놓은 것이라고 할 수 있다. 즉, 이동전화나 PDA와 같은 소형 무선 단말기 상에서 인터넷을 이용할 수 있는 기술을 말한다. 무선 데이터 시장과 인터넷은 매우 빠르게 성장하고 있으며, 계속해서 새로운 고객을 만들어 가고 있는 실정이다. 인터넷 사용자의 엄청난 증가와 인터넷의 폭발적인 성장은 새롭고 흥미 있는 정보 서비스의 창조를 요구한다. 현재까지 인터넷을 위해 개발된 대부분의 기술은 데스크탑이나 대형 컴퓨터, 그리고 중간이상의 높은 bandwidth, 일반적으로 신뢰할 수 있는 데이터 네트워크를 가져왔다. 반면에 Mass-market, handheld wireless device등의 이유로 wireless data network는 기존의 시스템에 비해 상당한 제약을 가진다. 대부분 현재의 mobile 네트워크는 사용자에게 제공될 수 있는 진보된 서비스를 포함하고 있다. 모바일 네트워크 오퍼레이터들은 모바일 네트워크 서비스의 사용을 촉진시키고 가입자의 이동 율을 감소시키기 위하여 유용하고 매력적인 방법으로 진보된 서비스를 제공하기 위해 노력하고 있다. WAP 스펙은 digital data networking standards, 그리고 IP, HTTP, XML, SSL, URLs와 같은 인터넷 기술들과 스크립팅과 다른 content format들과 같은 현재의 기술들을 확장한 스펙이라고 볼 수 있다.

2.2 I-Mode

WAP이 전 세계적인 무선 단말기 표준이라 하면 I-Mode는 일본의 NTT-Docomo에서 개발, 및 현재 가입자를 늘려가고 있는 새로운 기술이라 할 수 있다. I-Mode 전화기만으로 음성통화는 물론 은행구좌와 잔액조회, 이체를 비롯한 레스토랑 가이드, 타운 페이지 검색(각종 생활정보)등 생활에 밀접하고 편리한 온라인 서비스를 가능케 하는 서비스이다.

I-Mode의 특징은 종래의 음성 통화는 물론 I-Mode키를 누르기만 하면 여러 가지 서비스에 접속이 가능하다는 것이다. 제공되는 서비스는 Mobile Banking이나 티켓 예약 등 생활에 밀접하고 편리한 것을 중심으로 형성되어 있다. 또한 9.6Kbps의 패킷 통신이기 때문에 요금 산정 방식이 통화시간을 기준으로 한 것이 아니라 송수신한 데이터의 양으로 결정된다. I-Mode Mail은 I-Mode 전화기끼리는 물론 일반 E-Mail도 사용 가능하다. 사이트 접속 서비스는 때와 장소를 가리지 않고 원하는 정보 사이트로부터 조회가 가능하다. 예로, 외출 시나 이동 중에도 은행의 잔고조회나 각종 티켓 예약이 가능하며 휴대 전화기만으로는 인터넷에 접속이 가능하다. 현재는 텍스트와 일부 형식의 화상만 지원하고 I-Mode 형식으로 대응 가능한 홈페이지만 접속이 가능하다. 인터넷 접속 시는 URL의 입력이 필요하다. 메시지 서비스는 각 사이트에 접속해서 메시지 서비스를 신청하면 원하는 정보가 자동적으로 배달되는 서비스를 말하는데 뉴스나 날씨 정보 등을 받을 수 있다. I-Mode는 휴대전화기 간의 메일 송·수신은 물론이며 인터넷을 경유하여 E-Mail과의 송수신도 가능하다. 휴대전화번호가 그대로 E-Mail 주소로 쓰이게 되고 메일 주소의 변경도 가능하다. I-Mode의 구조를 보게되면 아래 그림과 같이 먼저 사이트에 접속하게 되고, 메일의 송·수신은 Docomo I-Mode 센터를 경유하여 이루어진다. I-Mode의 C-HTML은 HTML의 간소화된 형태로 몇 가지 Tag들이 사용되어진다.

2.3 ME(Mobile Explorer)

Stinger 프로젝트는 데이터와 음성 통신이 결합된 통신 단말기를 의미하며, PDA와 Phone 그리고 Pager가 하나의 단말기에 포함되어 다음과 같은 서비스의 제공을 목표로 한다.

- PIMS(Personal Information Management)
Pocket Outlook을 통해 연락처, 일정, 작업, 메일 정리 등이 가능하다.
- Internet Explorer를 탑재한다.
- 음성통신 : phone dialer, Call logging, 전화정보
- Office S/W 제공
윈도 CE에서 제공하던 Pocket워드와 액셀 viewer를 제공하고 Pocket Money와 Fax등의 사용이 가능하다.
- Game : 윈도 CE에서 제공하던 solitaire등의 게임이 가능하다.
- 이외에도 제어판에서 단말기의 기본적인 설정과 전화 설정 등의 기능을 제공한다.

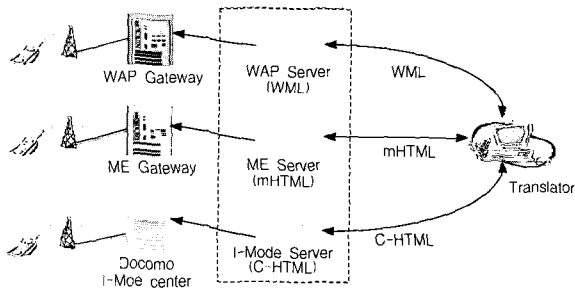
ME는 OS에 독립적이고 porting환경이 제공되므로 쉽게

porting이 가능하다. 사용되는 언어는 인터넷 표준인 HTML 3.2의 부분 집합인 M-HTML을 사용하고 이동전화기 및 PDA, 페이지 등의 소형기에 최적화되도록 설계하였다. 또한 느린 CPU와 한정된 메모리를 고려하여 복잡하지 않은 API들을 사용한다. ME는 WAP, HDML등과는 달리 이미 사용되고 있는 기존의 HTML 콘텐츠를 사용할 수 있도록 제공한다. 그러므로 WAP 및 HDML에 비해 C-HTML을 이용하는 I Mode와 같이 호환성 면에서는 앞서 있다. 또한 향후 인터넷 콘텐츠의 발전 방향을 XML로 보고 자사의 Windows Ce와 ME 역시 같은 방향으로 발전시키고 있다.[2-5, 8]

3. 시스템의 설계 및 구현

3.1 시스템 구조 설계

먼저 전체적인 시스템 구조를 설명하자면 아래 그림에서 보는 바와 같이 서버에 존재하는 각각의 콘텐츠를 변환기를 이용하여 구축하고자 하는 서버로 변환을 하게 된다. 변환을 할 때 모든 콘텐츠를 변환한 후에 에디터를 이용하여 콘텐츠 제공업자 측이 제공하고자 하는 서비스만을 재편집하여 WAP, I Mode, 또는 ME용 서버를 구축할 수 있다. 아래 그림에서는 세 가지 서비스를 제공할 수 있는 서버를 따로 두었지만 어차피 HTTP 기반에서 동작하기 때문에 하나의 서버에서 모든 서버를 동시에 연동시킬 수도 있다.



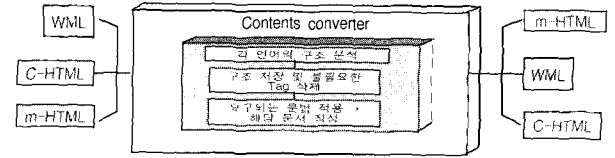
(그림 1) 시스템 전체 구조

세 가지 언어의 번역에 필요한 콘텐츠 변환기(Converter)에서는 다음과 같은 기능들이 필요하다. 변환기 내부에서 변환을 해주어야 할 각 소스마다 두 가지의 변환 결과를 내어야 하기 때문에 아래에 나열한 것처럼 6가지 경우의 변환 작업이 필요하다.

- C HTML to WML Converter
- mHTML to WML Converter
- WML to mHTML Converter
- C HTML to mHTML Converter
- WML to C HTML Converter
- mHTML to C HTML Converter

위에서 첫 번째 변환 작업을 예로 들면 C-HTML의 각 요소(element)에 대한 WML의 요소로서 대칭 되는 표를 만든다. 또한 각 요소에서 쓰여지는 속성들의 리스트 또한 표를 작성해서 각 요소들의 표와 속성들의 표를 이어주는 상속을 정

의한다. 상속되는 표에 의하여 분석되고 미리 저장되어진 구조에 맞게 요구되는 문법을 적용하여 문서를 작성하게 된다. 콘텐츠 변환기의 변환 과정을 정리하면 아래 그림과 같다.



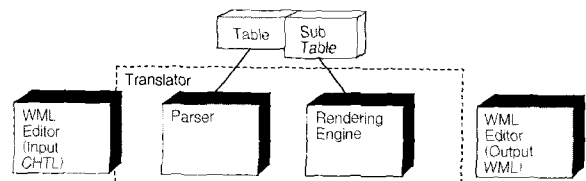
(그림 2) 콘텐츠 변환 원리

위와 같은 과정을 거쳐 콘텐츠 변환을 가진 문서를 저장하면 원문과 같은 기능을 하는 WML를 만들 수 있다. 콘텐츠 제공자가 원하는 서비스만을 제공하려면 에디터에서 다시 불러들여 편집을 한 다음 다시 저장을 하게 된다. 에디터에서는 클립 기능으로 태그를 추가할 때 해당 언어의 문법에 어긋나지 않도록 해주는 기능을 부여한다. 본 논문에서는 전체 시스템 구조를 아래와 같이 두 부분으로 나누고 비교적 단순한 구조인 C-HTML을 WML로 변환하기 위한 모듈의 구현을 목표로 한다.

- Contents Converter
 - : 세 가지 언어의 상호 호환을 위해 위에서 열거한 6가지 경우의 콘텐츠 변환이 필요.
- WML Editor
 - : 클립 기능을 이용하여 각 언어의 Tag 추가 및 콘텐츠 변경할 때 이용.[1, 7]

3.2 다중 언어 변환기의 내부 구조

본 논문에서는 앞서 열거한 6가지 경우의 콘텐츠 변환 중에 C HTML to WML의 변환을 예로 설명할 것이다. 변환기의 처리 순서는 아래 그림에서와 같이 먼저 파서를 이용하여 입력된 C HTML의 문서 구조를 분석한다. 각 내용을 분석한 후에 내용과 함께 분석된 정보를 저장한다. 구조가 파악된 후에는 모든 Tag들은 삭제되고 텍스트만 남게 된다. 이와 같은 처리가 끝난 후에는 저장된 정보와 각 요소에 미리 정의되어진 테이블을 이용하여 일반 텍스트에 WML 문법을 적용, WML 문서를 생성하게 된다. 생성된 문서는 WML Editor에서 콘텐츠 제공자에 의해 편집이 가능하도록 해준다.



(그림 3) 변환기 내부 구조

3.3 Rendering 과정

WML로의 변환에 앞서 WML 문서의 첫 머리에는 문서를 유효한 XML 문서임을 확인하기 위해 XML 선언과 DTD (Document Type Declaration)를 포함해야만 한다. 그러므로 먼저 다음과 같은 문서의 서두가 삽입되어야 한다.[6]

```
<?xml version = "1.0"?>
<!DOCTYPE wml PUBLIC "-//WAPFORUM//DTD WML 1.1
//EN" "http://www.wapforum.org/DTD/wml_1.1.xml">
```

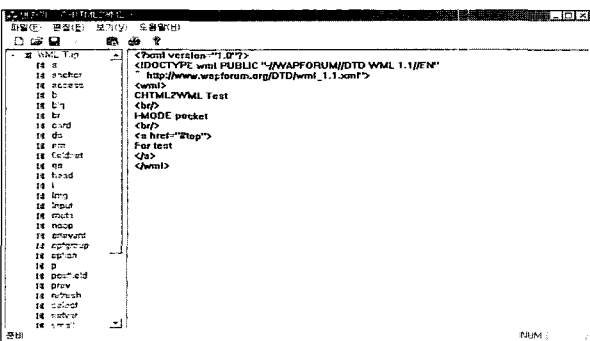
I-Mode에서 현재 서비스하기 위해 사용하는 C-HTML (501i)의 element에 대응하는 WML의 element를 비교, 정리한 몇 가지 예를 정리하면 아래 표와 같다. 나머지 경우는 지면상의 이유로 생략하였다.

| C-HTML Tag (501i, 502i에서 지원) | 설명 | 번호 | WML Tag |
|---|-------------------------|----|--|
| <!-- --> | Comments | 1 | <!- -> |
| <A> | Anchor | 2 | <go> <do> |
| accesskey attribute directkey function | Only C-HTML | 3 | Ignore |
| <BASE> | Base URL | 4 | <P> |
| <BLOCKQUOTE> | 인용문 사용 | 5 | <P> |
| <BODY> | 문서의 본문 테두리 | 6 | <card> |
| | 줄 바꿈 | 7 | |
| <CENTER> | 텍스트가 짧을 경우 엔 가운데로 정렬 | 8 | <P mode = "wrap" align = "center"/> |

C-HTML의 모든 element가 WML에 일대일 대응은 되지 않지만 accesskey를 제외하고는 비슷한 기능으로 대체할 수 있다. 만약 같은 기능을 하는 element가 없을 경우에는 WML에서 가장 비슷한 기능을 하도록 한다.

4. 결 론

본 논문에서는 세 가지 언어의 호환을 위한 변환기를 설계하고 구현하였다. 다중 언어 변환기를 개발하기 위해서 먼저 C-HTML 스펙과 WML 관련 스펙, 그리고 mHTML의 스펙을 분석하였고 각각의 서로 대응되는 요소들을 찾아 표를 작성하였다. 대응되는 요소들이 없는 경우에는 WML에서 지원하는 여러 가지의 태그들을 조합하여 최대한 같은 형식으로 디스플레이 해줄 수 있는 방안을 제시하였다. 변환기의 구조를 이해하기 위해서 설명한 C-HTML의 경우, HTML과 마찬가지로 문서 내에 내용을 정렬하기 위하여 몇 가지 Tag들을 이용한다. 이러한 것들은 WML의 P Tag와 그 속성들을 이용하여 정렬 방식이 같아지도록 하였다. C-HTML에서



(그림 4) 변환 후의 WML소스 출력

한다. 변환기를 실행 시키고 C-HTML to WML 변환 모듈을 추가된 기능중 하나인 전화 걸기 옵션은 WAP에서의 WTA (Wireless Telephony Application)에서 명시한 대로 변환을 선택한 후에 C-HTML의 구조를 가진 문서를 선택하면 아래 그림과 같이 WML로 변환된 문서의 소스를 출력한다.

편집기에서는 변환된 WML 소스를 콘텐츠 저작자가 클립 기능을 이용하여 다른 내용을 추가하거나 내용을 바꿀 수 있다. 변환기는 클립 기능을 추가 할 때에도 사용 언어의 문법에 맞지 않는 위치에 태그를 삽입할 수 없는 기능을 가지고 있다.

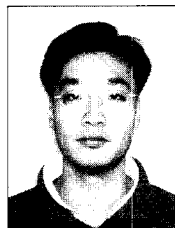
참 고 문 헌

- [1] Joe Sullivan, Tim Bickmore, Andreas Girgensohn, Digestor : Digesting HTML to HDML Feasibility Demonstration, Palo Alto, Aug. 20, 1997.
- [2] Dave Raggett, HTML 3.2 Reference Specification, W3C Working Draft, Sept 9, 1996.
- [3] Unwired Planet, HTML 2.0 Specification, Redwood Shores, June 15, 1997.
- [4] Tomihisa Kamada, Compact HTML For Small Information Appliances, February, 1998.
- [5] i-Mode : FAQ Frequently Asked Questions about i-Mode and the answers. February, 1998.
- [6] WAP WML Version 4-November-1999 , WAP Forum Specification, November, 1999.
- [7] Unwired Planet. WML1.1 and WML+1.1 to HDML Best-Effort-Translation.
- [8] http://www.nttdocomo.com/i/index.html.
- [9] Charles Arehart, Nirmal Chidambaram, etc, Professional WAP, WROX, 2000.



최 지원

e-mail : jackeroo@konkuk.ac.kr
 1999년 동국대학교 전자계산학과 학사
 2001년 건국대학교 컴퓨터공학과 석사
 2001년~현재 건국대학교 컴퓨터공학과 박사 과정
 관심분야 : 차세대 인터넷, 이동 컴퓨팅, WAP



김기천

e-mail : kckim@konkuk.ac.kr
 1988년 서울대학교 계산통계학과 학사
 1992년 미 Northwestern University 전산학 박사
 1992년~1996년 한국통신기술(주) 연구소 선임연구원

1996년~1998년 신세기통신(주) 기술연구소 책임연구원/차장
 1998년~현재 건국대학교 컴퓨터공학과 조교수
 관심분야 : 차세대 인터넷, 이동 컴퓨팅, IMT-2000