무선 인터넷 환경에서의 압축 기법을 이용한 다운로드 속도 개선에 관한 연구

오교환*, 정성우
고려대학교 컴퓨터정보통신대학원
hinick@korea.ac.kr*, swchung@korea.ac.kr

A study download performance enhancement Using compression method on wireless environment

KyoHwan Oh*, SungWoo Chung
Graduate School of Computer and Information Technology, Korea University

요 약

인터넷과 정보기기의 기술 발달로 개인 정보기기들이 소형화 집약화 되면서 이제는 어느 한 장소의 제한 된 공간이 아닌 자유롭게 이동하면서 통신을 할 수 있는 휴대인터넷 사용이 증가 하고 있다. 그러나 아직 휴대인터넷의 낮은 대역폭으로 인하여 사용자들이 만족할 만한 수준의 빠른 네트워크 속도를 제공 받지 못하는 것이 현실이다. 본 논문에서는 인터넷에서의 압축율이 좋은 파일에 압축기술을 적용하여 낮은 대역폭의 휴대인터넷에서 데이터 압축율에 따른 다운로드 시간 향상을 확인 해 보고자 한다.

1. 서 론

인터넷과 월드 와이드 웹(World Wide Web)의 등장으로 많은 사람들이 정보 기기를 이용하여 인터넷에 접속 하여 쉽고 빠르게 다양한 정보들을 이용할 수 있게 되었다. 초고속 인터넷의 발전으로 기존의 낮은 용량의 단순한 텍스트 위주의 웹 콘텐츠들이 시각적인 이미지와 동영상 위주의 고용량의 멀티미디어 콘텐츠로 변화 하였고, 정보기기의 고도화로 인하여 개인의 정보기기들이 소형화 집약화 되면서, 이제는 어느 한 장소의 제한 된 공간에서만 통신을 하는 것이 아닌 특정 공간에 구애 받지 않고 언제 어디서나 자유롭게 이동하면서 통신을 할 수 있는 휴대 인터넷 사용에 대한 관심이 높아 지게 되었다. 현재 상용화 되어 있는 휴대인터넷 서비스는 크게 HSDPA(High Speed Downlink Packet Access : 고속 하향 패킷 접속)와 Wibro(Wireless Broadband : 무선 휴대인터넷)로 나눌수 있다. HSDPA 서비스는 2GHz 대역의 주파수를 이용하고 있고, 주로 음성, 영상, 고속 데이터를 위한 비동기 방식의 이동 통신 서비스이다. HSDPA는 기존의 WCDMA보다 다운로드 속도를 향상시킨 3.5세대 이동통신으로 IP 멀티미디어 서비스 제공에 효율적이고 음성통신망에 기반을 두고 있으면서도 데이터 통신 다운로드 속도가 최고 14.4Mbps까지 가능한 서비스이다. Wibro는 Wireless와 Broadband의 합성어로서, 정지 및 이동 중에 언제 어디서나 높은 전송속도로 무선인터넷 접속이 가능한 서비스를 말한다. Wibro는 2.3GHz 대역을 이용하여 60Km/h 이상 이동시에도 무선 인터넷 서비스를 받을 수 있으며, 1Mbps급 이상의 전송속도로 이동 중에도 끊김 없는 인터넷 접속 환경을 제공 받을 수 있는 서비스이다.[1] 이와 같은 휴대인터넷 서비스의 실제 하향 최대 전송 속도는 HSDPA 서비스가 1Mbps, Wibro 서비스가 3Mbps로 유선인터넷의 100Mbps와 비교하면 무려 최대 100배의 속도 차이가 발생하게 된다.

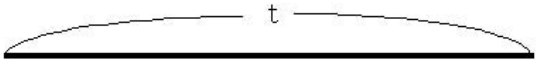
HTTP/1.1 protocol[2]을 적용하면 웹 자체의 content-encoding을 통하여 네트워크 대역폭을 절약할 수 있지만, 이는 자칫 압축에 따른 추가적인 오버헤드만 발생 시킬 수 있어 상황에 따른 압축을 할 수는 없다. 웹 콘텐츠들이 시각적인 이미지와 동영상 위주의 멀티미디어 콘텐츠로 변화하는 과정에서 자연스럽게 이와 관련 된 이미지와 동영상에 관련 된 압축 기술 또한 발전 하게 되었지만, 의외로 압축 효율성이 좋은 문서 파일 압축에 관해서는 지금까지 등한시 되어 왔던게 현실이다. 더불어, 정보기기 발전 속도에 비해 아직 휴대 인터넷에 대한 기술 발전이 빠르게 연동 되지 못하고 있는 시점에서, 압축 효율이 높은 문서 파일을 이용한 고도화 된 단말 장치들 간 낮은 대역폭 사이에서 어떻게 하면 좀 더 효율적으로 사용할 수 있을지를 연구해 보았다.

본 논문에서는 압축 효율이 좋은 파일들에 대하여 고도화 된 정보기기간 낮은 네트워크 대역폭 사이에서 효율적으로 이용할 수 있는 방법을 살펴본다. 2장에서는 관련 제안 방식을 알아보고, 3장에서는 관련 실험과 결과를 제시하고, 4장에서 결론을 맺는다.

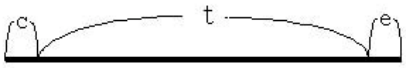
2. 제안 방식

네트워크 환경에서 이용하고자 하는 콘텐츠를 보다 빠르게 이용할 수 있는 방법은 첫째 높은 네트워크 대역폭을 이용하는 방법과 두번째 압축 효율을 높여 보다 작은 용량의 콘텐츠를 이용하는 방법이 있을 것이다. 그렇다면 낮은 대역폭의 휴대 인터넷 환경에서 이용하고자 하는 콘텐츠를 보다 빠르게 이용하기 위해서는 이용하고자 하는 콘텐츠의 압축율을 높여 보다 작은 사이즈로 원하는 콘텐츠를 보다 빠르게 다운로드 받아 이를 효과적으로 이용할 수 있을 것이다.

현단계의 가입자 평균 휴대인터넷 하향 최대 전송 속도는 HSDPA 서비스의 경우 1Mbps 이고, Wibro 서비스가 3Mbps 이지만, 실제 해당 서비스로 인터넷 속도측정 사이트 벤치비(http://www.benchbee.co.kr)를 통해 하향 속도를 측정한 결과 HSDPA 서비스는 0.45Mbps, Wibro 서비스는 2.42Mbps로 100Mbps로 서비스 되는 유선인터넷의 실 측정속도 83Mbps 보다 최대 185배의 차이를 보였다. 이렇듯 대역폭이 낮은 휴대 인터넷 환경에서 다운로드 속도를 개선 하고자 한다면 이용하고자 하는 콘텐츠의 용량을 압축 기법을 이용하여 일시적으로 용량을 줄인 다음 다운을 받고 다시 원래의 상태로 압축을 해제 한 후 이용 하면 될 것이다. (그림1)에서와 같이 압축 전과 후의 파일의 다운로드 속도를 비교해 보면 그 차이가 크다는 것을 알 수 있을 것이다. 즉, 압축 효율이 좋은 파일을 다운로드 할 경우 이를 서버에서 먼저 압축 하여 용량을 최소화 한 후 낮은 대역폭 사이에서 다운로드 받은 후 이를 다시 압축 해제 하여 이용하면 될 것이다.



(a) 일반 파일 다운로드 시간



(b) 압축 기술 적용 파일 다운로드 시간

(그림 1) 압축 기법을 이용한 다운로드 속도 개선

3. 실험

다운로드 시간 테스트의 휴대인터넷 서비스는 현재 전국적으로 상용 서비스 되고 있는 SKT의 HSDPA로 하였다. 표 1와 2는 성능 평가에서 사용한 기기들의 사양이다. 그리고 파일 다운로드 대상 서버는 웹호스팅 사이트(http://www.dothome.co.kr)에서 호스팅 서비스를 받아 테스트 파일을 서버로 업로드 하였고, 다운로드 시간 측정을 위해 Windows OS에 리눅스 환경 시뮬레이션 프로그램인 cygwin[3]을 설치 하였다.

일반적인 파일 다운로드 시간 측정과 파일 압축을 이용한 다운로드 시간 측정인 파일압축 시간 + 다운로드 시간 + 압축해제 시간을 각각 측정하여 비교 하였다.

<표 1> 테스트에 사용된 컴퓨터 사양

Computer	SX30-6B261/15R
CPU	Intel(R) CPU Pentium M 1500MHz
RAM	1GB
OS	Microsoft Windows XP
Utility	cygwin

<표 2> 테스트에 사용된 HSDPA 모뎀 사양

Company	SKT
MODEM	IM-H100
TYPE	NotePC 연결용 HSDPA USB 모뎀

3.1. 파일 압축

파일을 압축 하는 이유는 파일의 용량을 줄일 수 있고, 여러 개의 파일들을 하나의 파일로 만들어 간편히 보관, 전송, 관리 할 수 있기 때문이다. 압축의 기본 원리는 파일의 내용 중 반복되거나 필요 없는 정보를 제거하여 파일의 크기를 줄이는 것이다.

테스트 할 파일은 압축 효율이 좋은 문서 파일을 용량대 별로 임의선정 하였고, 압축이 이미 되어진 jpg 포멧도 압축율 비교를 위해 추가 하였다. 해당 파일의 압축은 무손실 압축 프로그램인 gzip[4]을 이용하여 압축 하였고, 용량을 최소화 하기 위해 gzip의 압축옵션 중 -9옵션을 이용하여 최대 압축을 하였다.

명령어 gzip -9 filename를 사용하여 표3에서와 같이 압축 전/후의 데이터 압축율을 비교 하였다.

<표 3> 파일 압축에 따른 데이터량 감소

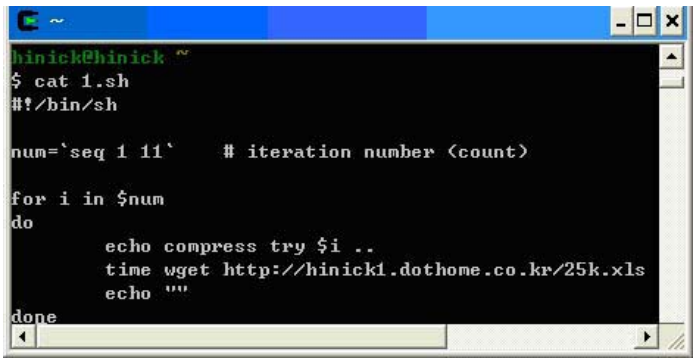
번호	파일 종류	Before Compression (bit)	After Compression (bit)	Percent Size (%)
1	XLS	25088	5683	77.3
2	PPT	51200	7760	84.8
3	DOC	101888	18654	81.7
4	DOC	255488	37230	85.4
5	PPT	512000	28383	94.5
6	XLS	1024000	766021	25.2
7	PDF	1968148	1665408	15.4
8	JPG	2046422	2043877	0.1
9	DOC	2047488	1660875	18.9
10	PPT	5162496	855216	83.4
11	DOC	10240000	6889689	32.7

임의로 선정한 문서 파일의 압축율을 테스트 한 결과 최대 94.5%의 높은 압축 효율이 있었고, 압축이 이미 되어진 jpg 포멧의 경우 중복 압축의 결과로 0.1%인 0%에 가까운 압축 효율이 발생함을 알 수 있었다.

3.2. 파일 다운로드 시간 측정

파일 다운로드는 wget[5] 프로그램을 이용하여 wget filename(웹에 링크되어 있는 파일의 절대경로) 명령어를 사용하여 다운로드 하였고, 시간 측정은 time 명령어를 이용하여 다운로드 시 소요 시간을 측정하였다. 시간 측정 방식은 1개의 파일당 10회에 걸쳐 시간을 측정 하였고, 서버의 응답시간 및 PC의 CPU 사용율 그리고 휴대인터넷의 네트워크 전송 속도 상황에 따라 약간의 차이가 발생하여 10회 측정한 결과의 평균값에서 근사한 5회만 가지고 재 평균을 내었다. 그림2 에서와 같이 파일 다운로드 스크립트를 이용하여 파일의 다운로드 시간을 측정을 해 보았다.

표4는 파일을 압축하지 않은 원본 상태의 파일을 다운로드 받을 때 걸리는 시간을 측정 한 결과이다.



(그림 2) 파일 다운로드 시간 측정 스크립트

<표 4> 압축 전 파일 다운 로드 시간

번호	파일 종류	1 차 (초)	2 차 (초)	3 차 (초)	4 차 (초)	5 차 (초)	평균 (초)
1	XLS	0.913	0.930	0.968	0.944	0.932	0.937
2	PPT	1.379	1.288	1.315	1.370	1.283	1.327
3	DOC	2.348	2.163	2.321	2.164	2.282	2.256
4	DOC	4.995	5.000	4.970	4.631	4.677	4.855
5	PPT	8.624	8.083	8.367	8.357	8.259	8.338
6	XLS	14.572	14.567	14.702	14.381	13.789	14.402
7	PDF	26.516	25.733	26.085	25.121	26.310	25.953
8	JPG	23.821	27.801	24.698	26.997	25.740	25.811
9	DOC	25.099	26.455	25.057	27.410	25.453	25.895
10	PPT	52.527	62.507	58.889	58.497	52.869	57.058
11	DOC	100.14	102.95	100.14	101.49	99.69	100.75

제안하고자 하는 압축율이 높은 파일을 다운 받기 전 서버에서 압축을 하고, 휴대인터넷을 통해 다운 받은 후 다시 다운받은 압축 파일을 압축해제 하는데 까지 걸리는 3단계의 시간을 각각 측정해 보았다. 여기서 첫번째 단계의 파일의 압축 시간에 대한 측정은 파일이 실제로 올려져 있는 웹서버에서 테스트 해야 하나, 웹서버를 컨트롤 할 수 있는 권한이 없기 때문에 통상적으로 서버의 하드웨어 사양이 일반 PC 보다 더 좋기 때문에 이에 걸리는 시간은 일반 PC보다 더 짧게 나을 수 있으나 동일하다는 가정하에 테스트PC에서 속도 측정을 하였다. 파일의 압축과 압축 해제 시간 측정 방식 또한 파일 다운로드 시간 측정과 같은 방법으로 측정하였다. 표5는 압축 효율이 좋은 문서파일에 압축기법을 적용 한 파일압축 시간, 다운로드 시간, 압축 해제 시간을 각각 측정 한 결과이다.

<표 5> 압축 기법 적용한 파일 다운 로드 시간

번호	파일 종류	압축 시간 (초)	다운 로드 (초)	해제 시간 (초)	최종 시간 (초)
1	XLS	0.062	0.491	0.057	0.610
2	PPT	0.107	0.661	0.062	0.830
3	DOC	0.081	0.823	0.060	0.964
4	DOC	0.109	1.123	0.063	1.295
5	PPT	0.173	0.997	0.065	1.235
6	XLS	0.228	9.225	0.082	9.535
7	PDF	0.318	17.910	0.100	18.328
8	JPG	0.335	22.426	0.101	22.862
9	DOC	0.439	18.707	0.116	19.262
10	PPT	1.002	9.725	0.138	10.865
11	DOC	1.423	71.083	0.262	72.768

<표 6> 압축 전/후 파일 다운 로드 시간

번호	파일 종류	Before time (초)	After time (초)	Percent time (%)
1	XLS	0.937	0.61	34.9
2	PPT	1.327	0.83	37.5
3	DOC	2.256	0.964	57.3
4	DOC	4.855	1.295	73.3
5	PPT	8.338	1.235	85.2
6	XLS	14.402	9.535	33.8
7	PDF	25.953	18.328	29.4
8	JPG	25.811	22.862	11.4
9	DOC	25.895	19.262	25.6
10	PPT	57.058	10.865	81.0
11	DOC	100.75	72.768	27.8

압축 전 파일 다운로드 시간과 압축기법을 이용한 파일 다운로드 시간을 비교 해 보았다. 표 6 에서와 같이 파일 압축효율에 따라 파일 다운로드 시간 효율도 비슷하게 나타나는 것을 알 수 있다.

4. 결론

본 논문에서는 현재 서비스 중인 휴대인터넷 HSDPA 에서 압축 효율이 좋은 문서 파일을 대상으로 압축 전과 압축 기술을 전용한 후의 다운로드 시간을 측정한 결과 압축 효율에 따라 최대 85%의 다운로드 시간을 단축 할 수 있었다. 파일의 압축율에 따라 틀리겠지만 테스트한 임의의 11 개의 파일의 평균 55%의 압축율과 45%의 시간 개선 효과를 보여 압축 효율이 좋은 파일에 압축기법을 적용하면 네트워크 대역폭을 그 전보다 크게 절약 할 수 있다는 것을 알 수 있었다. 이는 네트워크 대역폭을 절약한 비용 절감 및 효율적인 네트워크 사용이라는 결과를 가져 올 수 있을 것 같다.

참고문헌

[1] 김문구, 지경용, 박종현 “디지털 컨버전스 시대의 모바일 브로드밴드 전개:와이브로와 HSDPA” 한국전자통신연구원, 2006.

[2] The specification defines RFC2616 Hypertext Transfer Protocol - HTTP/1.1, http://www.w3.org

[3] The cygwin Linux-like environment for Windows utility, http://www.cygwin.com

[4] The gzip compression utility, http://www.gzip.org

[5] The wget retrieve files from the WWW utility, http://sunsite.dk/wget