

웹 로그 분석 모델

권영직*, 장광위**

요약 본 논문에서는 웹 관리자가 웹 사이트 분석을 위하여 웹 로그(web log) 분석을 통한 필요한 정보를 추출하고, 추출된 데이터를 분석하여 웹 환경을 개선시키고, 웹 사이트 방문자를 위해 더 좋은 서비스를 할 수 있는 새로운 웹 로그 분석 모델을 제안 한다.

이와 같은 목적을 달성하기 위해 본 논문에서는 웹 로그의 개념 및 유형, 웹 로그의 분석 원칙 및 웹 로그 분석 방법을 고찰하였다. 웹 로그 분석 도구로는 Web Log Expert를 활용하였다. 본 논문에서는 이 도구를 이용하여 웹 방문자 수에 대한 통계분석, 인기가 많은 웹 사이트에 대한 페이지분석 및 방문자 IP 분석 등을 실시하였다.

Key Words : 웹 로그 분석, Web Log Expert, 웹 방문자 수, 페이지 분석, IP 분석

I. 서론

월드 와이드 웹(WWW)의 발달로 다양한 분야에 대한 대량의 정보를 효과적으로 처리할 수 있게 되었고, 누구나 손쉽게 원하는 정보를 웹 검색을 통해 찾아낼 수 있게 되었다. 웹은 이러한 장점을 가지고 있지만, 너무 많은 웹 사용자로 하여금 검색을 통해 얻어진 방대한 양의 결과물들로부터 필요한 정보를 선별하기 위한 시간과 비용을 낭비하게 되는 단점도 있다. 따라서 이러한 문제를 해결하기 위한 연구의 필요성이 점차 증가하였는데, 그 중 웹 로그 파일 분석은 웹 이용자들의 웹 사이트에 대한 웹 로그 데이터를 분석하고, 그 데이터를 이용하여 동적으로 변화하는 환경에 적응하기 위하여 맞춤형 웹 사이트 서비스를 제공하려는 것이 대표적인 연구의 사례이다.

따라서 본 논문에서는 이와 같은 추세에 맞추어 웹 로그(web log) 분석을 위한 필요한 데이터를 추출하고, 추출된 데이터를 분

석하여 웹 환경을 개선시키고, 웹 사이트 방문자를 위해 더 좋은 서비스를 제공할 수 있는 새로운 웹 로그 분석 모델을 제안 한다. 웹 로그 분석 방법에는 가장 일반적인 분석 방법이 ①경로 분석, ②연관 규칙, ③순차 패턴, ④분류 및 클러스터링이 있다.[5] 본 논문에서는 이들 4가지 방법 중에서 통계학적 관점을 이용하여 비교적 방문자 수가 많은 웹 사이트 4개를 대상으로 하여 실험을 한 후에 가장 효과적인 웹 로그 분석 모델을 제안 한다.

II. 웹 로그 분석에 관한 이론적 배경

2.1. 웹 로그의 개념 및 유형

2.1.1 웹 로그의 개념

“웹 로그”란 웹 사용자들이 웹 사이트를 방문하고 제품이나 서비스를 구매하는 과정을 통해 발생하는 데이터이다. 정보획득이나 구매를 목적으로 웹 사이트를 방문하는 방문자들은 웹 로그의 형태로 웹 사이트 내에 흔적을 남기는데 이러한 데이터를 기반으로 해서 다양한 정보를 추출해 내는 것이 웹 로그 분석이다(Cooly, Mobasher and Srivastava, 1997).[4]

* 대구대학교 컴퓨터·IT공학부 교수

** 대구대학교 대학원 컴퓨터정보공학과 석사과정

웹 로그 파일(web log file)은 웹 서버(web sever)에서 발생하는 모든 데이터들을 일정한 시간 단위로 기록하고 있는 것이다. 웹 로그 파일 분석은 이러한 웹 로그를 이용해 웹 사이트를 방문한 사람들의 방문 데이터를 웹 서버가 인식하여, 누가, 언제, 무엇을 요청했고 가져갔는지, 웹 서버에 얼마나 많은 사용자가 접속하였는지, 어디에서 얼마나 머무렀는지, 어떠한 페이지를 자주 방문하였는지, 가장 오래 보는 또는 가장 많이 보는 페이지는 무엇인지 등의 방문자 기록이 정리된 파일을 분석하는 방법이다(Borges & Levene, 1999).[4]

웹 로그 파일에 저장되어 있는 주요 내용은 <표 1>과 같다.

② Extended Log Format:

“01:31:35 210.182.144.220 GET /Emp/Guide/t042.html 200”

저장된 로그 데이터는 웹 서버에 접속한 방문객의 IP(Internet Protocol) 주소, 접근 시간, 접근 방법, 대상 URL, 전송 프로토콜, 에러 코드, 전송 바이트 수와 같이 방문객을 인식할 수 있는 정보와 웹 페이지에 대한 방문 정보들을 포함하고 있으며, 다음과 같은 특징을 가지고 있다.[5]

- 첫째, 데이터를 수집하기 위한 특별한 절차가 필요 없다.
- 둘째, 분석에 불필요한 정보를 많이 포함하고 있다.

Field	Description	Data format
Date	Date, time and time-zone of request	Date
Time taken	Time taken for transaction to complete	Time
User_ IP	Remote host IP or DNS entry	String
User name	Remote login name of the user	String
Server	Server name, IP address and port	String
Bytes	Bytes transferred, sent or received	Int
Request	URL query and stem	URL
Status	HTTP status code returned to the client	Int
Protocol version	Used transfer protocol and its version	String
Referrer	Previous page	URL
User agent	Service provider	String
System	Client operating system	String
Cookie	Cookie ID	String

<표 1> 웹 로그 파일의 주요 내용

2.1.2 웹 로그의 유형

“웹 로그 데이터”는 웹 서버에 모든 방문객들의 접근을 기록한 데이터로 아래와 같이 HTTP 프로토콜의 일부로 명시된 Common Log Format 또는 Extended Log Format을 따라 저장된다.

① Common Log Format:

“66. 249. 65. 144-anonymous [25/Feb/2007: 04: 02: 35 +0800]“GET /dzzsread. php? id=2331HTTP/1. 1” 200 11493”

셋째, 분석을 위해서는 추가적인 정보가 필요하다.

넷째, 일반적으로 대용량이다.

2.2 웹 로그의 분석원칙

웹 서버에는 아래 <그림 1>과 같이 정보 받기, 처리 요청 및 실행 오류 등 여러 가지 기준 정보를 기록한다. 웹 관리자는 웹 서버 상에서 웹 로그의 통계, 웹 로그 분석을 통해 웹 서버의 효율성 및 효과성을 발견하고, 웹 서버의 잘못된 원인을 배제할 뿐만 아니

라, 사용자의 방문 특성과 복원(분배) 법칙에 대한 경우를 이해하여 웹 서버 상에서 더 좋은 기능을 제공해야 한다. 웹 서비스 모델에는 다음과 같이 세 가지 단계가 있다.[3]

① 서비스 요청 : 웹 브라우저를 통해 웹 사용자는 웹 서버에 필요한 데이터 요청을 발송하고 웹 서버는 HTTP 프로토콜에 따라 웹 사용자에게 관련되는 많은 기본 정보를 송부한다. ② 서비스 응답 : 웹 서버는 웹 사용자의 요구에 대응하는 기능을 실행하고, 필요한 정보를 웹 사용자에게 반환 요청을 받는다. 오류가 발생하면 오류 코드가 보내진다.

③ 로그에 추가 : 웹 서버는 웹 사용자가 방문 과정 중에 관련 정보에 대한 추가 정보가 있으면 로그 파일에 추가하여 저장한다.

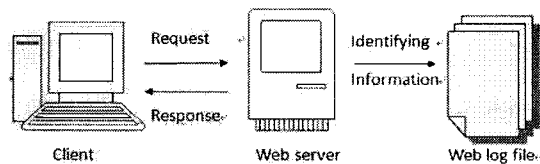


그림 1. 웹 서비스 모델

2.3 웹 로그 분석

2.3.1 웹 로그 분석 시 고려 요인

웹 로그 데이터는 일반적으로 대용량인 특징이 있다. 대용량 데이터를 분석할 때 이용할 수 있는 방법은 여러 가지가 있을 수 있겠지만, 아래와 같이 3가지만 고찰하여 두었다.[5]

- ① 통계적 관점: 표본을 추출하여 이용하는 방법
- ② 컴퓨팅 관점: 병렬 또는 분산 컴퓨팅 방법;
- ③ 데이터베이스 활용 관점: 요약 정보를 갖는 테이블을 이용하는 방법 등이 있다.

2.3.2 웹 로그의 분석 방법

① 방문 경로 분석

사용자 방문 경로(user web navigation path)는 웹 방문자가 한 행동을 구체화 한 것으로, 사용자 방문 경로는 웹 로그의 전체

처리과정을 통해 추출되는 세션(session) 정보에서 얻을 수 있다. 사용자 웹 방문 패스는 페이지 명(page name=url), 페이지에 머문 시간(duration)으로 구성되어 진다.[4]

② 연관 규칙

일반적으로 로그 데이터 분석에서 연관 규칙의 유용성을 측정하는 기준은 지원(support)과 신뢰도(confidence)이다. 예를 들면, 사용자(A)가 페이지 /E-business/M2를 방문하는 동시에 페이지 /E-business/M7를 방문했고, 사용자(B)는 페이지 /E-business/M3를 방문하는 동시에 페이지 /E-business/M7를 방문했다고 가정하면, /E-business/M7은 A와 B가 공동으로 방문한 페이지가 된다. 따라서 이것은 A와 B 사이에 일정한 관계가 있다고 할 수 있다. 이런 연관규칙을 이용하여 좀 더 좋은 웹 사이트를 구성할 수 있다.

③ 순차 패턴

순차 패턴은 발견은 웹 페이지를 방문하는 순서 정보를 이용하여 몇 개의 대표적인 패턴을 찾기 위한 방법으로 웹 사이트의 효율적 구성에 도움을 주며, 그 결과는 다른 분석, 예를들면 “분류 등”에도 이용될 수 있다. 순차 패턴에 관한 알고리즘 및 세부적인 내용은 Agrawal 과 Srikant(1995), Pei(2000)을 참조하기 바란다.[5]

④ 분류 및 클러스터링

클러스터링은 웹 페이지 또는 방문자들을 몇 개의 군으로 나누어 웹 사이트의 특성을 파악하기 위한 방법으로 방문자 관리를 위해 가장 많이 활용되는 기법이다.

2.3.3 웹 로그 분석을 위한 측정 요인

① 히트 수(Hits): 히트 수는 방문자가 웹 사이트를 접속했을 때 연결된 파일의 숫자를 합하는 것으로 한 페이지를 전송할 때 그 안에 포함된 그래픽, HTML등의 파일을 히트로 계산한다.

② 방문자수(Visiter): 방문자 수는 정확

이 말해서, 실제로 방문한 사람 수는 아니다. 웹 사이트 내에서 일정시간 동안 지속적인 움직임이 있었던 활동을 하나의 단위로 하여 그 수를 측정한 것이다.

③ 페이지 뷰(Page View): 페이지 뷰는 하나의 HTML문서를 보는 것을 말하며 인터넷 사용자가 웹 사이트를 접속할 때 본 페이지를 페이지 뷰로 계산한다.

④ 대역폭(Bandwidth): 대역폭은 인터넷 회선의 속도로서 웹 사이트에 여러 명이 접속 하면 할수록 좀 더 빠르고 넓은 대역폭의 통신 회선이 필요 하다.

⑤ 웹 페이지 분석: 많이 찾는 페이지와 많이 찾는 디렉토리에 대해 정보를 분석한다.

⑥ 방문자 IP 분석: 최근방문자 IP분석이나, 지역, 회사/기관별 IP 분석 등이 있다.

웹 페이지 분석 및 방문자 IP 분석은 웹 로그 파일에서 추출한 데이터(히트수, 방문자수, 페이지 뷰 등)를 이용하여 분석하고 그 결과는 표나 그림으로 만들어 제시하게 된다.

III. 본 논문에서의 연구 모형

위에서 고찰한 이론적 내용을 토대로 웹 로그 분석을 위한 연구의 모형을 아래 <그림 2>와 같이 도출하였다.

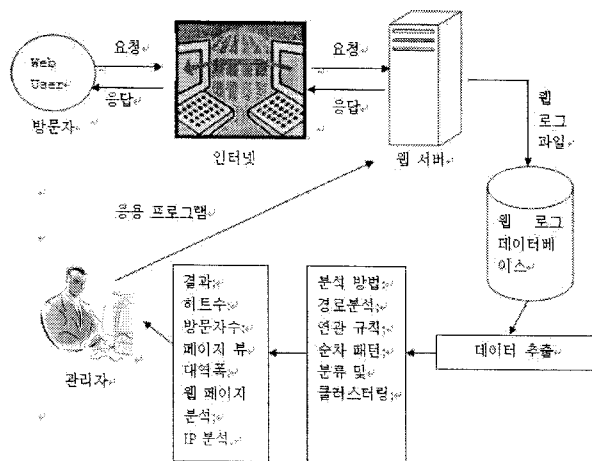


그림 2 본 논문에서의 연구 모형

IV. 실험

웹 로그 분석에 사용한 데이터는 인기 있는 쇼핑몰 사이트인 <http://www.smsync.com>, www.taobao.com, www.paipai.com, www.eachnet.com에서 추출하고, 실험도구로는 “Web Log Expert Lite”를 이용하였다. “Web Log Expert Lite”는 Web 서버의 다양한 로그 파일에 대하여 종합분석하는 전문적인 작은 소프트웨어 도구이다. 이 소프트웨어는 히트수, 방문자수, 페이지 뷰, 대역폭, 웹 페이지 분석, IP분석 등을 할 수 있다. 이 이외에도 웹 로그 파일 중에서, 방문 시간, 가장 인기 있는 페이지 순위, 오류 통계, 일반적인 접근 경로, 브라우저 및 플랫폼 통계, 사용자 플랫폼 등을 분석할 수 있다. 이와 같은 로그 파일 통계 분석 정보를 통해 웹 사이트 관리자는 웹 사이트의 전체적인 상황에 대한 숫자화 또는 정확한 인식을 통해 웹 사이트 설계 및 최적화와 웹 사이트 내용의 개선 등을 통해 웹 사이트 사용자를 위한 더 좋은 서비스를 제공할 수 있다.

V. 실험 결과 분석

5.1 웹 로그에 의한 기본 분석

인기 있는 쇼핑몰 사이트 중에서 <http://www.smsync.com>를 대상으로 웹 로그에 의한 기본적인 분석 데이터로는 일별 총 히트 수, 방문자수, 페이지 뷰, 대역폭을 이용하였다. 이들에 대한 웹 로그 분석 결과는 <표 2>과 같다.

총 히트수	5.546
페이지 뷰	934
총 방문자수	931
총 대역폭	187.16MB

<표 2> 웹 로그에 의한 기본적인 분석 결과

위의 <표 2>에서 방문한 방문객의 총 히트 수는 5,546번이며, 총 페이지뷰는 934 page이다. 또한 방문객은 931명이고, 총 대역 폭은 187.16MB이다.

이런 방법으로 중국의 인기 있는 C to C 쇼핑몰 사이트에 대하여 실험을 한 결과

<표 3> 같은 결과를 얻었다.

마찬가지로 beauty.taobao.com/의 가장 인기 있는 페이지의 히트수는 “13.2”%, paipai.com/의 히트수는 “17.1”%,

쇼핑몰 사이트	히트수	방문자수	페이지뷰	대역폭
http://www.taobao.com/	59,349,973	8,563,214	22,258,606	2,140.80GB
http://www.paipai.com/	31,023,670	6,293,578	15,851,376	1,567.52GB
http://www.eachnet.com/	30,273,054	5,369,871	17,109,613	1,863.56GB

<표 3> 3개의 쇼핑몰 사이트에 대한 웹 로그 분석

동일한 시간내에 taobao 사이트는 방문한 방문객의 총 히트 수는 59,349,973번이며 총 페이지뷰는 22,258,606 page이고, 방문객은 8,563,214명에 달하는 것을 볼 수 있다. 또한 대역폭은 제일 커서 중국 쇼핑몰 사이트 분야 중에서 제1순위에 있다. 마찬가지로 “paipai”, “eachnet” 사이트들도 위와 같은 방법으로 평가될 수 있고, 이들 쇼핑몰 사이트들을 비교해 보면 paipai 사이트는 eachnet 사이트보다 히트수와 방문자수에 있어서는 우수하지만, 페이지 뷰와 대역폭에서는 열등하다. 따라서 paipai 사이트는 두 번째로 큰 C to C 쇼핑몰 사이트이지만 페이지 뷰와 대역폭에 있어서 보완이 요망된다.

5.2 웹 로그에 의한 가장 인기 있는 방문 페이지 분석

.eachnet.com의 히트수는 “12.9”%이다. 그 가운데 요청이 최고 많은 <http://www.smsync.com/image/>의 히트수가 3.270점에 도달해서 이 웹 사이트 페이지의 총 히트수의 67%이상을 넘고 방문자수가 총 방문자수의 56.9% 보유한다. 마찬가지로 쇼핑몰 사이트들에 대하여 taobao 네트워크의 화장품의 페이지가 인기 제일 많고 히트수의 비율과 방문자수의 비율은 13.2%, 15.6% 이다. paipai 네트워크에서 사용자들은 여성용품에게 관심이 제일 많고 히트수의 비율과 방문자수의 비율은 17.1%, 14.8% 이다. eachnet 네트워크는 전자상품을 제일 잘 팔고 히트수의 비율과 방문자수의 비율은 12.9%, 13.5% 이다. 사용자가 최고 관심하는 페이지로서 관리자는 반드시 이 웹 페이지가 잘 틀 할뿐만 아니라 최저의 시간에 웹 페이지를 열수 있도록 보증해야한다.

예를 들어 설명하면 A html는 방문한 웹 사이트 홈 페이지 이고 Z html는 웹 사이트

웹 사이트	히트 수	비율	방문자 수	비율
http://www.smsync.com/images/	3,270	67.1%	530	56.9%
http://beauty.taobao.com/	7,834,196	13.2%	1,335,861	15.6%
http://www.paipai.com/lady/	5,305,048	17.1%	931,450	14.8%
http://www.eachnet.com/zone/digital/	3,905,224	12.9%	724,933	13.5%

<표 4> 웹 사이트에서 가장 인기 있는 방문 페이지 데이터

위의 <표 4>는 4개의 웹 사이트 중에서 가장 인기 있는 방문 페이지를 측정한 데이터이다. 위의 <표 4>에서 “smsync.com 사이트”의 가장 인기 있는 페이지의 히트수는 “3,270”이고, <표 2>에서 총 히트수가 “5.546”이므로 비율은 67.1%에 달한다.

제일 인기 많은 페이지 이다. 홈 페이지부터 Z html를 찾을 과정을 분석하면 웹 페이지 요청 과정의 페이지에 머문 시간 감소 및 웹 페이지 방문 경로 선택에 대하여 문제들을 처리 방법을 고려해야한다. <그림 4>는 웹 페이지 요청 과정 이다. 조사 따라서 사용자가 웹 페이지 방문할 때 1초이상 시간에 기

다리면 불편해진다. A html에서 Z html를 열 때에 사진과 그림, flash, 문서 등을 초기 다운로드 하는 시간 필요 있고 페이지에 머문 시간도 있다. 그래서 초기 다운로드 필요 없는 내용을 제거하고 대역폭도 일정정도 넓혀 준다. 그리고 웹 사이트 관리자는 Z html에서 광고를 증가를 동시에 웹 페이지를 요청 시간을 연장할지 여부를 주목해야한다. A html부터 Z html까지 방문 경로를 정확한 선택이 중요하다. 웹 페이지를 방문한 경로는 여러 개 있는데 <그림 5>가 제시한다. ②경로는 방문 과정이 너무 복잡해서 찾으려면 시간이 많이 지난다. ③경로는 ②경로보다 더 낫다. Z html는 사이트의 제일 유행한 페이지 때문에 많은 방문자들이 쉽게 찾기를 위하여 홈 페이지에서 바로 보일 수 있으면 제일 좋다. 그래서 ①경로는 제일 좋은 경로이다.

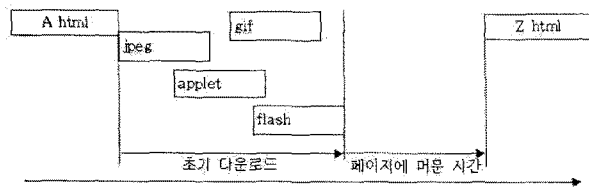


그림4 웹 페이지 요청 과정

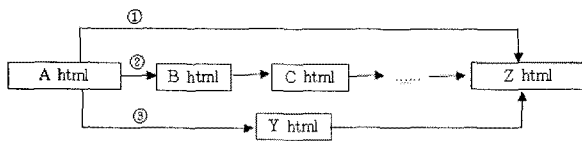


그림5 웹 페이지 방문 경로

5.3 웹 로그 의한 방문자 IP분석

Top Hosts			
Host	Visits	Requests	Bandwidth (KB)
1 202.300.182.12	54	26	752
2 62.202.26	31	8	8
3 24.48.117.22	16	14	12
4 208.203.152.0/8	26	13	56
5 192.243.179	11	11	2
6 66.77.28.11	10	10	2
7 217.70.229.44	42	5	128
8 81.138.8.133	10	5	248
9 198.187.140.14	13	4	3,115
10 198.51.216.66	3	4	27
11 213.84.73.92	1	4	12
12 202.247.48.28	3	1	23
13 63.108.178.11	2	1	13
14 66.77.24.27	3	4	36
15 208.247.401.94	3	4	24
16 208.174.201.88	24	3	203
17 24.213.151.93	5	1	2
18 66.12.26.72	2	3	27
19 194.108.237.35	238	2	737
20 84.491.51.3	3	3	146
21 195.252.49.240	28	3	1,201
22 198.138.94.202	3	3	39
23 62.301.2	14	3	12
24 208.208.46.208	3	3	13
25 66.94.219.87	44	3	2,257
26 208.247.401.140	3	3	39
27 211.156.47.13	15	15	12
28 202.183.0.37	15	3	7
29 202.154.26.212	16	2	105
30 217.39.233.7	1	1	11

표5 방문자의 IP 주소 표(예)

<표 5>는 <http://www.smsync.com>에 대한 방문자들의 전 30개의 IP주소와 연관된 정보를 얻을 수 있다. Web Log Expert Lite가 로그세션의 데이터통계에 의해 우리는 그 가운데서 방문자가 최고 들어있는 30개 IP주소를 추출했다 한 개의 특징적인 IP데이터베이스를 형성하고 이미 알려져 있는 세계의 IP 데이터베이스와 비교를 진행 한다. <그림 6>는 방문자의 IP정보를 이용하여 방문자 IP 분석 과정 흐름 것이다.

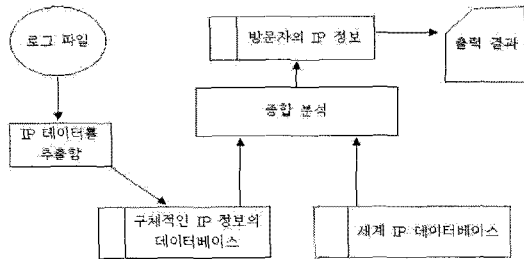


그림6 방문자 IP 분석 흐름

방문자의 IP유통

미국 유렵 아시아 기타 지역

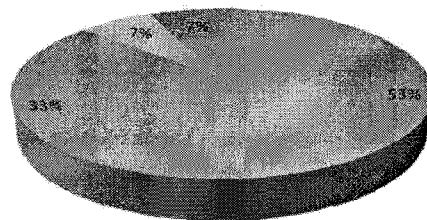


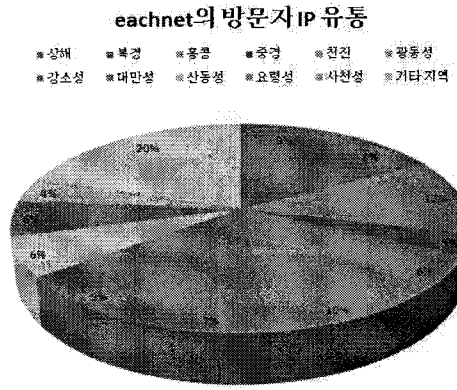
그림7 방문자의 IP유통

<그림 7> 에서 분명하게 볼 수 있다 싶은 IP를 방문하는 것은 어느 지역에서 온 것이다. 미국의 사용자가 세계총사용자의 절반이상을 차지한다. 아시아와 기타 지역의 사용자 방문수량은 비교적 작다. 그래서 웹 사이트의 페이지디자인과 풍경이 주로 미국과 유럽사용자의 사용습관에 따라서 개선을 진행한다 뿐만 아니라 이 두 지역의 정보교체량도 넓혔다.

<http://www.smsync.com>에 대하여 분석 과정을 따라서 <그림 8>는 중국의 세계 쇼 팡몰 사이트의 방문자 IP 지역별 분석 결과

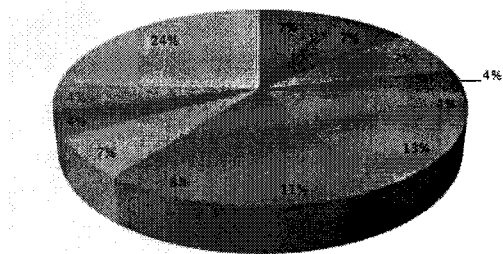
를 나올 것이다. taobao 네트워크의 방문자들이 주로 광둥성, 강소성, 대만, 상해, 북경 등 경제발달과 인구 미집한 지역에서 집중한다. 그런데 paipai 네트워크의 방문자 분포는 상대한 평균적인 상황을 있다. 엄청난 경쟁 압력 때문에 새로 생긴 paipai 네트워크는 더 많은 사용자 유치를 위해 중국에서 두 번째 - 계층 도시에 경영 전략적 이동한다. eachnet 네트워크는 국제적인 네트워크인데 사용자들이 홍콩, 광둥, 상해, 대만처럼 국제화 도시에서 많이 있다.

eachnet의 방문자 IP 유동



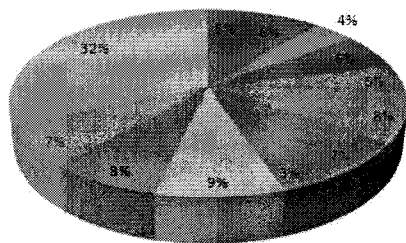
이상에 분석 결과들은 향후에 아래와 같은 인터넷 마케팅을 하는데 유용한 자료로 활용될 수 있다.

- ① 교차 판매 및 계획 수립
- ② 캠페인 효과에 대한 평가
- ③ 개인화된 서비스 등을



제공할 수 있는 기본 자료로 활용할 수 있기 때문에 로그 데이터의 적절할 분석과 활용이 필수 요소로 부각되고 있다.

paipai의 방문자 IP 유동



6 결론과 상후 연구

결론: 본 논문은 web log의 연구와 web log file의 분석을 진행하여 하나의 비교적 표준적인 web log분석흐름모델을 제출했다. 모델에 따라 web log 분석의 과정에 대하여 알 수 있다. web log의 통계공구를 사용하여 웹사이트의 방문 데이터 분석을 진행하여 웹사이트의 히트수, 방문자수, 방문자 IP 등을 얻을 수 있으므로 ①접속 유휴 시간 감소 ②방문자에 대한 정보 분석 ③웹사이트 재구축 전략 평가 등 결과를 얻었다. 웹 사이트 관리자를 위해 더 편한 관리를 제공해준다.

그림 8 taobao, paipai, eachnet의 방문자 IP 유동 분석

상후 연구: QoS (Quality of Service, 서비스 품질)기술은 다음 - 세대에 네트워크 핵심 기술 중 하나이다. 이 기술은 컴퓨터 네트워크 및 멀티미디어 통신 연구 및 초점 문제를 개발하고, 특정 대역폭, 지연에 대한 정량, 지연 지터, 손실 속도, 처리량 및 기타 성능 지표가 있다. 인기가 많은 사이트에 기하 급수적인 성장의 결과는 인터넷으로 HTTP 요청을 종종 서버 과부하의 문제에 직면하게 된다.[2]

참고 문헌

- [1] H. Cenk Ozmutlu, Amanda Spink, Analysis of large data logs: an application of Poisson sampling on excite web queries, Information Processing and Management 38 (2002) 473-490.
- [2] Tian Changpeng, Based on the Analysising and Researching Web server Log of Web QoS, Chongqing Technology and Business University, 400067.
- [3] LIU Zuoda, XU Jingfang, CHEN Maoke, LIXing, Campus Network User Analysis Method Based on Web Log Mining, (1. Department of Electronic Engineering qinghua University, 2. Network Research Center of qinghua University, Beijing100084, China).
- [4] 김정하, 김소은, 웹 로그 분석을 통한 여행사 웹 방문 경로에 관한 연구, 관광·레저연구, 제16권, 제3호(통권 제28호) pp. 209 ~ 226, 한국관광·레저학회 2004. 9.
- [5] 김석기, 안정용, 한경수, 웹 로그(Web Log) 데이터 분석 방법에 관한 연구, 응용통계연구 제14권 2호, 2001년, pp. 261-271.
- [6] 정선경, 이철우, 웹 로그 분석을 적용한 웹 사이트내의 웹 컨텐츠 분석 연구, 2003년도 한국정보과학회 춘계 학술발표논문집 Vol.30, No. 1. 579-581.
- [7] 김영철, 강명구, 김기수, 웹 로그 분석 방식의 시스템 확장성 연구, 전자상거래학회지, 2008년. 제9권, 제2호, pp. 3-19.
- [8]
<http://naver.com/sally542/100021121212>