

# 동적 웹 페이지의 로그 분석에 관한 연구

김완규\*, 김현주\*, 김 일\*\*

## 요약

포털, 쇼핑몰, 인터넷 신문 등의 웹사이트는 접속 시간이나 접속자에 따라 동일한 주소에 접속해도 다른 콘텐츠를 제공하고 있다. 기존의 클릭스트림 로그 분석 방법은 접속자마다 다른 정보를 제공하는 개인화 시스템과 롤오버 탭 기능을 사용하여 접속자의 반응에 따라 다른 정보를 노출하는 동적 웹페이지에 대한 로그 분석이 불가능하다. 기존의 데이터 수집 및 분석 방법은 페이지 뷰, Hits 알고리즘 등을 이용한 빈도의 성향을 기본으로 하고 있다. 이러한 방법은 웹 브라우저에서 발생하는 단순 클릭 이벤트를 중심으로 분석하기 때문에 롤오버 탭 기능이 있는 동적 웹페이지에서는 롤오버 탭 기능의 사용에 대한 분석을 할 수 없는 한계를 가지고 있다. 본 논문에서는 웹 브라우저의 스크롤바 위치와 윈도우 커서의 움직임을 주기적으로 수집하여 로그 분석 서버로 전송하고, 로그 분석 서버는 해당 웹페이지에 대한 시각적 인식영역과 동적 웹페이지 요소를 분석하여 캡처한 이미지와 수집한 정보를 분석하는 동적 웹페이지 로그 분석 시스템을 설계하고 구현하였다.

## A Study on Log Analysis of Dynamic Web Pages

Wan-Kyoo Kim\*, Hyun-Ju Kim\*, Il Kim\*\*

## ABSTRACT

Web sites such as portals, shopping malls, and online newspapers provide users with different contents depending on time or person of accessing, even though the users access the same URL. It is impossible for the existing log analysis method of click-stream to conduct log analyses in the following two aspects: first, in the personalization system which provides different information for each user; second, in the dynamic web pages on which different contents are exposed depending on the responses of the user by using roll-over tab functions. In this paper, we seek to investigate the dynamic web pages using roll-over tab functions. We, by doing so, design and implement the log analysis system which analyzes the exposure and response of contents provided in the dynamic web pages.

Key Words : dynamic web page, click-stream, personalization, log analysis system

---

\* 경남과학기술대학교 컴퓨터융합공학과(✉ wankyoo@gntech.ac.kr)

\*\* S-네트워크 연구원

· 제1저자(First Author) : 김완규 · 교신저자(Correspondent Author) : 김 일

· 접수일(2012년 1월 16일), 수정일(1차 : 2012년 2월 13일), 게재확정일(2012년 2월 16일)

1. 서론

용이 노출되는 기능을 롤오버 탭이라 한다.

인터넷 웹사이트를 통해 제공되는 동적 웹페이지는 동일한 웹페이지에 접속하여도 접속 사용자, 접속 시기 등에 따라 다른 콘텐츠가 노출되는 개인화 시스템에서 많이 사용되었다[2,9,10,11].

개인화 시스템을 구현하기 위해서는 동적으로 제공되는 웹페이지에 대한 로그 분석이 필수적이다. 동적 웹페이지에 대한 로그 분석은 초기에는 웹페이지의 파라미터에 값을 데이터베이스에 저장하고 이를 매핑하며 분석하는 방법을 사용하였다.

그러나 롤오버 탭(roll-over tab) 등의 기능을 사용하여 접속자의 반응에 따라 다른 콘텐츠를 노출하는 동적 웹페이지에 대한 로그 분석은 불가능하다[4].

본 논문에서는 롤오버 탭 기능을 이용한 동적 웹페이지에 대하여 고찰하며, 동적웹페이지에서 제공되는 콘텐츠의 노출과 반응을 분석하는 로그 분석 시스템을 설계하고 구현하였다.

본 논문의 구성은 다음과 같다. 먼저 2장의 관련 연구에서는 롤오버 탭을 이용한 UI에 관하여 살펴보고, 3장에서는 본 논문에서 제안하는 페이지 스크롤과 윈도우 커서 움직임을 이용한 동적 웹페이지 로그 분석 시스템을 설계한다. 4장에서는 동적 웹페이지 로그 분석 시스템을 구현 및 평가한 후, 5장에서는 결론 및 향후 연구 과제를 제시한다.

II. 관련 연구

2.1 롤오버 탭 UI(User Interface)

그림 1은 롤오버 탭을 이용한 동적 웹페이지의 예를 설명한 그림이다. 롤오버 탭이란 처음 접속했을 때 사용자에게 노출되는 내용이 1과 같이 A영역인 뉴스면일 경우, 사용자가 경제면에 관심이 있어 B영역으로 마우스를 오버하거나 클릭하면 2와 같이 경제면의 내



그림 1. 롤오버 탭  
Fig. 1 Roll-over tab

롤오버 탭을 사용함으로써 웹페이지는 제한된 영역에 사용자의 반응을 이용하여 웹페이지를 재 호출하지 않고 많은 내용을 노출시킬 수 있는 장점을 가진다.

그림 2는 포털사이트에 대한 롤오버 탭 기능을 조사한 것으로, 롤오버 탭 기능을 사용한 곳은 총 9곳으로 약 90개의 내용이 비노출 속성으로 다운로드 되고 사용자의 반응에 따라 내용이 노출된다.

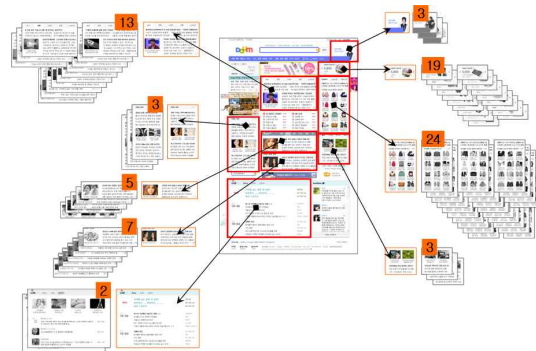


그림 2. 동적 웹페이지 예(www.daum.net)  
Fig. 2 Dynamic web pages

표 1은 동적 웹페이지를 사용하고 있는 대표적인 웹 사이트의 메인화면에서 동적 요소인 롤오버 탭을 사용한 요소수와 롤오버 탭에 의해 노출되는 면적을 픽셀단위로 분석하였다.

표. 1 동적 웹페이지의 롤오버 탭 분석  
Table. 1 Analysis of roll-over tab on dynamic web pages

순번	사이트	메인화면 크기	롤오버 탭 요소	비노출 요소 개수	롤오버 탭 면적 비율	비노출 면적	비노출 비율
1	www.daum.net	1004x1587	9	90	45%	6,367,206	338%
2	www.naver.com	880x1319	8	133	60%	17,072,860	1471%
3	www.nate.com	982x1549	16	77	40%	5,238,214	344%
4	www.yahoo.co.kr	972x1413	9	91	32%	6,283,846	458%
5	www.nexon.com	1064x925	4	23	33%	2,441,368	248%
6	www.hangame.com	967x1568	4	9	28%	810,624	54%
7	www.netmarble.net	981x1380	6	19	31%	1,241,781	92%
8	www.chosun.com	970x2747	10	31	16%	1,261,435	47%
9	www.joinmsn.com	927x1334	9	47	58%	4,506,280	348%
10	www.auction.co.kr	993x1984	13	52	47%	3,180,930	161%
11	www.11st.co.kr	986x2599	21	174	71%	12,657,880	494%

대부분의 웹사이트가 약 880픽셀에서 1064픽셀 사이의 너비를 사용하여 모니터의 해상도가 1024x768 픽셀 이상일 경우 탐색에 불편함이 없도록 디자인되었다.

롤오버 탭 요소는 작게는 4개소에서 많게는 21개소까지 차이가 많았으며 롤오버 탭 요소에 반응하는 비노출 내용의 개수도 9개에서 174개까지 그 차이가 컸다.

롤오버 탭을 사용한 레이아웃의 면적은 작게는 16%에서 많게는 71%까지 사용되고 있으며, 대부분의 웹페이지가 전체 레이아웃의 30% 이상을 롤오버 탭으로 사용되는 것으로 조사되었다.

특히, 비노출 면적의 비율은 메인 화면 크기 대비 작게는 47%에서 많게는 1471%까지 차지하는 것으로 조

사되어 롤오버 탭과 같이 동적 웹페이지에 대한 선호도가 높은 것으로 조사되었다.

동적 웹페이지의 사용은 2008년 이후 그 사용이 급격히 증가하고 있으며, 미국의 경우 최근 리뉴얼되는 웹사이트의 대부분이 동적 웹페이지를 사용하고 있는 것으로 조사되었다.

### III. 동적 웹페이지 로그 분석 시스템

#### 3.1 롤오버 탭 태깅

본 논문에서 제안하는 동적 웹페이지 로그 분석 시스템에서 롤오버 탭의 기능 검출과 롤오버 탭 이용 로그 정보를 생성하기 위해서는 먼저 롤오버 탭을 사용한 HTML 코드에 대한 분석과 로그 태깅 방법을 설명한다.

그림 1의 롤오버 탭을 이용하는 코드는 그림 3과 같다.

```

/* css 코드 시작 */
#media div.daumhide{display:none;}
#media .box{zoom:1;}
/* css 코드 끝 */
<!-- 미디어다음(중복) 영역 시작 -->
<div class="news" id="media">
<h2 class="cmtab"><span class="daumhide">미디어다음</span></h2>
<h3 class="cmtab_tit" id="newsTitle">
<a href="news=title">뉴스</a></h3>
<div class="box" id="news">
..... </div>
<h3 class="cmtab_tit" id="economyTitle">
<a href="economy=title">경제</a></h3>
<div class="box daumhide" id="economy">
..... </div>
.....</div>
<!-- 미디어다음(중복) 영역 끝 -->
    
```

그림 3. 롤오버 탭 구현 HTML 코드 예  
Fig. 3. HTML codes for implementation of roll-over tab

그림 3의 HTML 코드는 그림 1의 웹페이지에서 롤오버 탭 영역을 표시하기 위한 프로그램 코드로 그림 1의 영역 A에서 영역 E까지는 마우스오버 이벤트가 발

생하면 각 영역에 해당하는 레이어를 보이거나 숨기는 기능을 수행하게 된다.

여기서, 영역 A를 나타내는 “뉴스” (<h3 class="cmtab\_tit" id="newsTitle"><a href="news.html">뉴스 </a></h3>)에 마우스가 오버되면 OnMouseOver 이벤트가 발생하여 CSS 코드의 div.daumhide가 해제되면서 뉴스 영역(<div class="box" id="news">)이 화면에 표시되며 각각의 id에 따라 표시되는 내용은 표 2와 같다.

표. 2 롤오버 탭을 위한 ID  
Table. 2 IDs for roll-over tab

OnMouseOver id	newsTitle	economyTitle	sportsTitle	entertainTitle	hotIssueTitle
display id	news	economy	sports	entertain	hotIssue

한편, 그림 3의 프로그램 코드에서 id가 사용된 엘리먼트의 위치와 크기를 분석하면 표 3과 같다.

표. 3 롤오버 탭 ID 분석  
Table. 3 Analysis of IDs for roll-over tab

id 이름	left	top	width	height	id 포함 여부	중복 여부	마우스오버 영역 이름
media	321	284	400	248	o	x	
newsTitle	321	284	80	32	x	x	A
economyTitle	401	284	80	32	x	x	B
sportsTitle	481	284	80	32	x	x	C
entertainTitle	561	284	80	32	x	x	D
hotIssueTitle	641	284	80	32	x	x	E
news	321	316	400	207	x	o	
economy	321	316	400	207	x	o	
sports	321	316	400	207	x	o	
entertain	321	316	400	207	x	o	
hotIssue	321	316	400	207	x	o	

표 3에서 media는 내부에 id를 사용한 엘리먼트를 포함하고 있고, “news”, “economy”, “sports”, “entertain”, “hotIssue” 가 동일한 영역을 사용하고 있는 id임을 알 수 있다.

그리고 표 3과 같이 동일한 영역을 사용하는 id의 부모 id인 “media” 를 롤오버 탭 영역의 이름으로 정의하고, 롤오버 탭 영역은 (321, 316, 400, 207)이다.

확인된 롤오버 탭 영역을 구분하기 위하여 동일한 롤오버 탭 영역을 사용하는 id의 엘리먼트를 상위 개념의 단어로 묶어서 구분하고, 이를 부모 id로 정의하며, 부모 id에는 하위 엘리먼트들이 공통으로 사용하는 DIV의 좌표를 기록하면 된다.

본 논문에서는 롤오버 탭 영역을 검출하고 이용 현황을 로그 정보로 생성하기 위하여 표 4와 같이 표준대칭 규칙을 제안한다.

“OMP\_media” 에서 “OMP\_” 는 롤오버 탭 영역을 나타내는 엘리먼트를 표시하는 id이고 “OMC\_news” 에서 “OMC\_” 는 롤오버되는 내용의 엘리먼트를 표시하는 것으로 “OMP\_” 로 시작하는 id를 분석하면 롤오버 탭 영역의 검출이 가능하며, “OMP\_” 영역에 노출되는 엘리먼트를 확인하기 위해서는 웹브라우저가 웹페이지를 접속하였을 때 자바스크립트로 “OMP\_” 영역에 노출되는 “OMC\_” 엘리먼트를 찾아 로그 정보를 생성하고 로그 서버에 전송함으로써 동적 웹페이지의 롤오버 탭에 대한 로그 분석이 가능하다.

표. 4 동적웹페이지 로그분석을 위한 롤오버 탭 ID수정

Table. 4 Modified IDs of roll-over tab for log analysis

기존 id	media	news	economy	sports	entertain	hotIssue
수정 id	OMP_media	OMC_news	OMC_economy	OMC_sports	OMC_entertain	OMC_hotIssue



표. 6 페이지 접속 로그 정보 전송 예  
Table. 6 Transmission of log information

```
url=/Interest/49.html,referrer=http://www.waurank.com,time=2010-6-15/13-57-7.858,zone=-9,sw=1280,sh=1024,bw=1272,bh=915,color=16,cookie=y,java=y,history=0
```

두 번째 단계는 이용자의 이용 행태 정보를 지정된 샘플링 간격으로 수집하여 전송한다.

이용자 행태 정보를 수신한 이용 정보 분석기는 수신된 정보를 분석하여 이용 데이터베이스에 저장한다.

표 7은 수집한 이용자 행태 정보를 나타내며, 표 8은 전송된 이용자 행태 정보를 나타낸다.

이용자 행태 정보 수집에 사용된 샘플링 간격은 0.1 초를 사용했다.

0.1초 마다 수집되는 이용자 행태 정보를 이용 정보 분석기로 정보를 전송할 경우, 상당한 트래픽이 발생되어 샘플링 데이터를 10개 또는 20개로 패키징하여 전송한다[4,6,8,11].

표. 7 이용자 행태 정보  
Table. 7 Information of user type

구분	데이터형	데이터 예
URL ID	정수형	1
샘플링 ID	정수형	1
웹브라우저 스크롤 위치	날짜형	240
윈도우 커서 위치(X)	정수형	449
윈도우 커서 위치(Y)	정수형	340

표. 8 이용자 행태 정보 전송 예  
Table. 8 Transmission of user type information

```
uri=0&MT=T150X449Y340T250X449Y340T350X449Y340T450X449Y340T550X449Y340T650X449Y340T750X449Y340T850X449Y340T950X449Y340T1050X449Y340
```

표 9는 전송할 롤오버 탭 상태 정보를 나타내며, 표 10은 전송된 로그 데이터를 나타낸다.

표. 9 웹페이지 접속 정보  
Table. 9 Access information on web pages

구분	데이터형	데이터 예
eid	스트링	sociallink
nu	정수형	1
user	스트링	eyerank
url	날짜형	/Interest/49.html
id_1	스트링	OMP_media
pid	스트링	
t_1	스트링	미디어
x_1	정수형	356
y_1	정수형	38.5
w_1	정수형	350
h_1	스트링	43

표. 10 롤오버 탭 사용 로그 정보 전송 예  
Table. 10 Transmission of log information on roll-over tab

```
RO.asp?eid=sociallink&nu=1&user=eyerank&url=/Interest/49.html&id_1=OMP_media&pid=&t_1=%C6%BC%Cl%F6%C4%DC&x_1=356&y_1=38.5&w_1=350&h_1=43&id_2=OMP_news&pid=OMP_media&t_2=%C6%BC%Cl%F6%C4%DC&x_2=803&y_2=252&w_2=225&h_2=46&idcnt=2
```

### 3.3 동적 웹페이지 로그 분석 시스템

본 논문에서 제안하는 동적 웹페이지 로그 분석 시스템은 그림 6과 같이 구성된다.

그림 6에서 UAC(User Activity Collector)는 웹페이지에 접속한 사용자의 움직임(Activity)을 수집하여 로그 분석 서버로 전송하고 로그 분석 서버는 웹사이트에 접속한 사용자의 행위 정보를 수신하여 동적 웹페이지에 대한 로그 분석을 수행한다.

로그 분석 서버는 웹페이지를 시각적으로 분석하고 웹페이지에 포함된 롤오버 탭을 탐지하는 기능과 UAC로부터 수집된 사용자의 움직임 정보를 분석하여 동적 웹페이지에 대한 로그 분석을 수행한다. 웹페이지를 시각적으로 분석하기 위한 프로세스는 다음과 같이 진행된다.

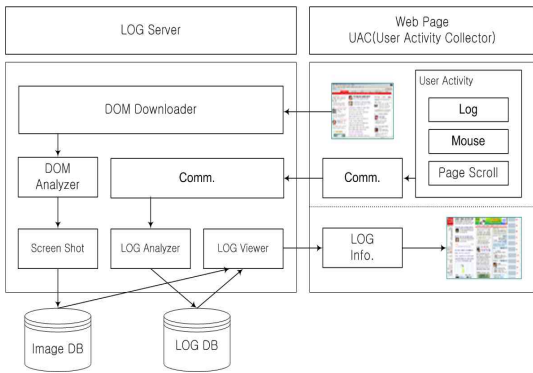


그림. 6 동적 웹페이지 로그 분석 시스템 구성도  
Fig. 6 System diagram of log analysis of dynamic web pages

먼저 웹페이지를 다운로드하여 DOM(Document Object Model) 형태로 시각적으로 표현한 후 DOM 분석기에 의해 웹페이지에 포함된 웹 콘텐츠의 시각적인 표현 위치 및 크기 정보를 픽셀 단위로 분석한다. 화면 캡처를 통하여 웹화면 및 영역을 이미지로 생성한 후 웹페이지에는 UAC를 삽입한다.

웹 브라우저에 다운로드된 UAC는 웹페이지 정보 및 윈도우 커서 정보와 페이지 스크롤 정보를 포함한 사용자의 움직임 정보를 주기적으로 수집하여 통신모듈을 통하여 로그 분석 서버로 전송한다.

로그 분석 서버의 LOG Analyzer는 UAC로부터 전송되어온 웹페이지 로그 정보와 사용자의 행위 정보를 분석하여 데이터베이스에 저장한다. LOG Visualizer는 동적 웹페이지의 롤오버 탭 요소에 대한 시각적 분석 정보와 사용자 행위 정보들을 시각화하여 웹페이지 형태로 제공한다.

#### IV. 구현 및 평가

본 논문에서 제안한 동적 웹페이지 로그 분석 시스템의 사용자 활동 정보의 수집 시간 간격은 L. Granka

의 눈시선 추적 분석[11]에 사용된 최적 시간인 0.2초~0.3초의 1/2인 0.1초를 적용하였으며, 구현 환경은 표 10과 같다[3,6].

표. 10 H/W 환경  
Table. 10 Hardware environments

구분	운영 체제	CPU	메모리	하드디스크	용도
H/W #1	Windows 2003 Server	4-Core Xeon 2.4GHz×1	4GB	250GB×2 (RAID 1)	DB 서버 (MSSQL 2008)
H/W #2	Windows 2003 Server	4-Core Xeon 2.4GHz×1	2GB	250GB×1	웹 서버 (IIS 6.0)
H/W #3	Windows XP Professional	Dual Core 2.0GHz×1	2GB	320GB	클라이언트

표. 11 주요 시스템별 사용 언어 및 H/W  
Table. 11 Languages and hardwares used

구분	사용 언어	사용 H/W
DOM Analyzer	VC++ 6.0	H/W #1, H/W #3
UAC	JavaScript	H/W #3
LOG Analyzer	VC++ 6.0, ASP	H/W #1, H/W #2
LOG Viewer	ASP	H/W #1, H/W #2

본 논문에서는 웹페이지에 대한 시각적인 인식 영역을 분석하고 웹브라우저의 스크롤바 위치 및 윈도우 커서의 위치를 수집하는 시스템과 이를 분석하는 동적 웹페이지 로그 분석 시스템을 설계하고 구현하였다.

그림 7과 같이 동적 웹페이지를 생성하고 2곳의 영역에 롤오버 탭 기능을 사용하였으며, 왼쪽의 “공지사항”에 2개, “POPUP ZONE”에 5개의 내용이 롤오버 탭으로 사용자의 반응에 따라 다른 내용이 노출되도록 하였다.

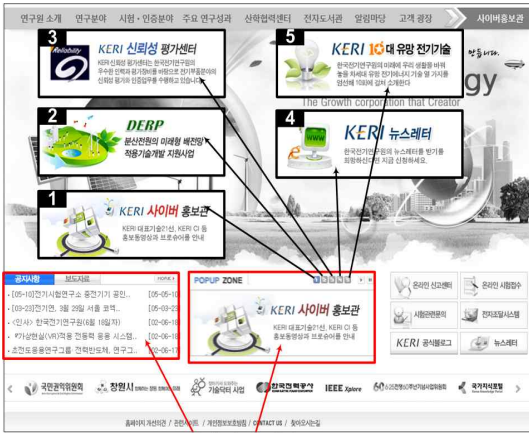


그림 7. 동적 웹페이지  
Fig. 7 Dynamic web pages

그림 8은 동적 웹페이지에 포함된 롤오버 탭 영역에 대하여 노출 시간 분석 및 윈도우 커서 활동 분석을 통한 로그 분석 결과를 나타낸 것이다.



그림 8. 롤오버 탭 영역 분석  
Fig. 8 Analysis of roll-over tab area

그림 9는 동적 웹페이지에 대하여 노출 시간 분석 및 윈도우 커서 활동 분석을 통한 웹페이지에 대한 히트맵 분석 결과를 나타낸 것으로 적색은 활동이 왕성한 영역이며 청색은 활동이 거의 없는 영역이다.

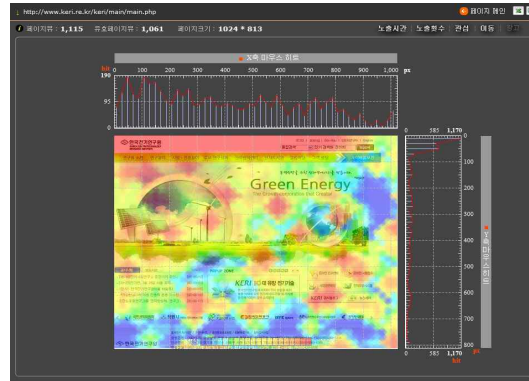


그림 9. 윈도우 커서 활동 정보를 이용한 히트맵 분석  
Fig. 9 Analysis of hit map

그림 10과 그림 11은 웹페이지에 접속한 사용자에 대하여 사용자 행동 정보를 수집하고 이를 세로 픽셀 당 노출 횟수와 노출 시간을 초 단위로 분석한 것이다.

그림 10에서 웹브라우저에 표현된 처음 화면에 포함된 영역의 노출 횟수가 많고 화면 아래쪽으로 내려갈수록 노출 횟수가 점점 감소함을 알 수 있다.



그림 10. 위치별 노출 횟수  
Fig. 10 The number of times being exposed at each position

웹페이지의 노출 횟수 및 노출 시간은 제공되는 콘텐츠와 레이아웃에 따라 상이하게 나타나므로 페이지

스크롤에 의한 사용자 활동 정보를 분석한 결과로 웹 페이지에 대한 상세 노출 분석이 가능하여 웹페이지 중 어느 부분이 사용자에게 가장 많이 노출되었는지 쉽게 분석할 수 있다.



그림 11. 위치별 노출 시간  
Fig. 11 Time being exposed at each position

그림 12는 웹페이지에 접속한 이용자의 행위 정보를 수집하고 이를 표현한 것으로 이용자의 마우스 활동 데이터는 동적 웹페이지에 대한 로그 분석에 있어서 가장 중요한 데이터로 개인별 행위 정보를 나타낸다.

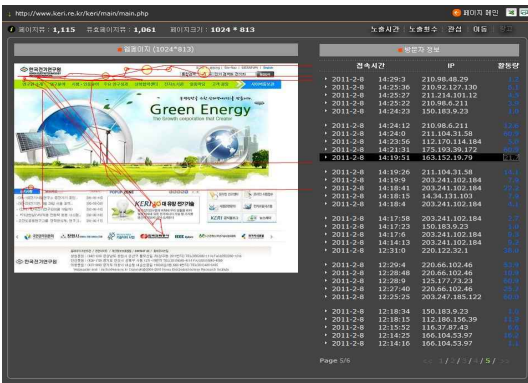


그림 12. 접속 사용자별 마우스 커서 움직임 정보  
Fig. 12 Movement Information of mouse cursor by each user

## V. 결론

본 논문에서는 웹 브라우저의 스크롤바 위치와 윈도우 커서의 움직임에 주기적으로 수집하여 로그 분석 서버로 전송하고, 로그 분석 서버는 해당 웹페이지에 대한 시각적 인식영역과 동적 웹페이지 요소를 분석하여 캡처한 이미지와 수집한 정보를 분석하는 동적 웹페이지 로그 분석 시스템을 설계하고 분석하였다. 기존의 데이터 수집 및 분석 방법은 페이지 뷰, Hits 알고리즘 등을 이용한 빈도의 성향을 기본으로 하고 있다. 이러한 방법은 웹 브라우저에서 발생하는 단순 클릭 이벤트를 중심으로 분석하기 때문에 롤오버 탭 기능이 있는 동적 웹페이지에서는 롤오버 탭 기능의 사용에 대한 분석을 할 수 없는 한계를 가지고 있다.

본 논문에서 제안한 시스템은 기존 분석 방법의 한계를 극복하기 위해 웹브라우저 내에서 사용자의 움직임을 찾아내어 웹페이지에 대한 관심도를 측정하였다.

본 논문에서는 웹 사용 마이닝에 사용되지 않았던 윈도우 커서의 좌표정보 및 긴 페이지의 스크롤 정보를 활용하여, 동적 웹페이지에 접속한 사용자가 어떤 정보에 관심을 나타내는지 정확하게 분석할 수 있는 방법을 제시하였다.

본 논문에서 제안한 시스템은 최근 사용이 급증하고 있는 그림 7의 롤오버 탭 기능을 포함한 웹페이지의 사용을 확대하여 스마트폰과 같이 작은 화면을 가진 장치에서 여러 번 접속하지 않고도 콘텐츠 서비스를 가능하게 함으로써 사용자의 편의성을 증대시킬 수 있다.

사용자 편의성 증대와 함께 그림 8의 롤오버 탭 기능에 대한 로그 분석을 가능하게 함으로써 웹페이지의 클릭수와 체류 시간으로 관심도를 분석하는 기존 방법을 극복하여 롤오버 영역 탭 기능에 대한 반응 시간 및 클릭수로 관심도 분석 범위를 확대하였다.

향후 연구 방향으로서는 본 논문에서 제안한 로그 분석 시스템을 인터넷 광고, 인터넷 쇼핑몰 등에 적용하는 태깅 방법 및 노출 분석 방법에 대한 다양한 연구와 함께 사용자가 접속했던 동일한 환경과 동일한 움직임 보여줄 수 있는 뷰어에 대한 연구가 필요하다.

### 참고문헌

- [1] Yang, T. and Zhang, H., "HTML Page Analysis Based on Visual Cues", In 6th International Conference on Document Analysis and Recognition(ICDAR 2001), Seattle, Washington, USA, 2001.
- [2] Chang Doo Lee, Bong Jun Choi, Zoo Gang, Il Kim, Kyoo Seok Park, "A Study on Detection Technique of Bookmarking using Favicon in Web Browser", Proceedings of International Conference on East-Asian Language Processing and Internet Information Technology 2002, EALPIIT2002, Korea Information Processing Society, pp.427-433, 2002.01.
- [3] Milos Kovacevic, Michelangelo Diligenti, Marco Gori, Marco Maggini, Veljko Milutinovic, "Reconition of Common Areas in a Web page Using Visual Information: a possible application in a page classification", Proceedings of the 2002 IEEE International Conference on Data Mining (ICDM 2002), pp. 250-257, 2002
- [4] Youn-Hong Jung, Il Kim, Kyoo-Seok Park, "Design and Implementation of an Interestingness Analysis System for Web Personalization & Customization", Journal of Korea Multimedia Society, Vol. 6, No. 4, July 2003, 707-713.
- [5] D. Cai, S. Yu, J. -R. Wen, and W. -Y. Ma, VIPS:"a vision-based page segmentation algorithm", Microsoft Technical Report, MSR\_TR-2003-79, 2003.
- [6] L. Granka, T. Joachims, and G. Gay, "Eye-Tracking Analysis of User Behavior in WWW-Search", Poster Abstract, Proceedings of the Conference on Research and Development in Information Retrieval (SIGIR), 2004.
- [7] Bong Joon Choi, Il Kim, Yong Won Shin, Kwang Hyung Kim, Kyoo Seok Park:"A Study of Link-Competition in a Hyperlinked Environment". International Conference on Internet Computing 2004: 339-344
- [8] Il Kim, KyooSeok Park, "Vision Based Web Usage Mining", Proceedings of The 2005 International Symposium on Multimedia Applications in Edutainment(MAEDU2005), pp.18-21, 2005.
- [9] 최봉준, 김일, 김원규, 박규석, "웹 브라우저의 Favicon을 이용한 즐겨찾기 등록 탐지 시스템", 한국정보과학회 제9회 학술발표논문집 제9회 1호, pp.93-98, 2001
- [10] 김일, 최봉준, 박규석, "웹브라우저의 즐겨찾기 행동에 대한 탐지기술 연구", 경남대학교정보통신연구소 연구논문집 제 4권, pp.47-61, 2002.12.
- [11] 김일, 박규석, "웹컨텐츠 마이닝을 이용한 개인화 시스템의 설계 및 구현", 경남대학교공업기술연구소 연구논문집 제21 권 1호, pp.5-14, 2003.03.
- [12] 김일, 박규석, "하이퍼링크 관리를 위한 시각화 엔진의 설계 및 구현" 연구논문집, 경남대학교정보통신연구소, 2003.12.31
- [13] 김일, 박규석, "하이퍼링크 환경에서의 링크 분석 시스템 설계 및 구현", 경남대학교 대학원논문집 제19집 제2권 통권27권, pp.215-226, 2004.

### 감사의 글

이 논문은 2010학년도 경남과학기술대학교 기성회 연구비에 의하여 연구되었음

### 저자소개



김원규(Wan-Kyoo Kim)

1985년 경남대학교 전자계산학과(공학사)  
 1988년 경남대학교 산업대학원 컴퓨터공학과(공학석사)  
 1998년 경남대학교 대학원 컴퓨터공학과(공학박사)  
 1989년~현재 경남과학기술대학교 컴퓨터융합공학과 교수  
 ※ 관심분야: 운영체제, 멀티미디어시스템, 인터넷응용



김현주(Hyun-Joo Kim)

1988년 경상대학교 전산통계(이학사)  
1990년 숭실대학교 대학원 전자계산(공학석사)  
2000년 경상대학교 대학원 컴퓨터과학과(이학박사)

2002년~현재 경남과학기술대학교 컴퓨터융합공학과 부교수

※ 관심분야: 정보검색, XML, 컴퓨터교육, 데이터마이닝

김일(IL Kim)



1997년 경남대학교 전자계산학과(공학사)  
1999년 경남대학교 대학원 컴퓨터공학과(공학석사)  
2006년 경남대학교 대학원 컴퓨터공학과(공학박사)

2010년~현재 S네트워크 연구원

※ 관심분야: 데이터마이닝, 정보검색, 전자상거래