

# 웹 모니터링 시스템 구현 및 활용 방안 연구\*

## : 쇼핑몰 이용고객에 대한 응답소요시간 평가를 중심으로

박철제\*\*·조재균\*\*\*

### 〈 목 차 〉

I. 서론	3.3 리포팅
II. 관련연구	IV. 효과분석 및 평가
2.1 에이전트	V. 결 론
2.2 웹 변경 탐지	참고문헌
III. 시스템 설계 및 구현	Abstract
3.1 에이전트	
3.2 인덱싱	

## I. 서론

인터넷 쇼핑몰 시장의 급속한 성장과 더불어 쇼핑몰간의 경쟁도 심화되고 있다. 고객의 입장에서는 쇼핑몰간의 경쟁이 심화될수록 가격이나 품질 만족도가 높은 상품을 선택할 기회가 많아지면서 고객서비스에 대한 불평행동을 적극적으로 하게 되었다. 한편, 이러한 고객의 불평행동은 기업에 엄청난 파급효과를 미치게 되었으며 각 기업에서도 이에 적극적으로 대응하여 쇼핑몰 이용고객에 대한 응답소요시간을 향상시키고자 많은 노력을 하고 있다. 그러나 인

터넷을 통한 정보교류의 확산성 및 정보과다 (information overload)로 인하여 고객의 불평행동에 관한 정보의 획득이 힘들고 적절한 시간 내에 대응하지 못함으로써 응답소요시간을 향상시키지 못하는 문제가 있다. 실제 각 기업에서는 고객센터 운영에 많은 비용을 지불하고 있으나 웹 상의 고객응대 측면에서는 만족할 만한 성과를 거두지 못하고 있다. 이는 인터넷의 활성화로 유통되는 정보의 양이 폭발적으로 증가하여 여러 웹사이트를 방문하고 수작업으로 그 목록을 관리하는데 한계에 봉착하고 있기 때문이다.

\* 본 논문은 2005년도 동의대학교 교내연구비(과제번호: 2005AA073)에 의해 조성되었음

\*\* 동의대학교 경영정보학과 부교수, cjpark@deu.ac.kr

\*\*\* 동의대학교 e비즈니스학과 부교수, jgcho@deu.ac.kr

이러한 문제를 해결하기 위한 방안으로서 에이전트 시스템을 이용한 정보검색이 일반화되고 있다. 그러나 정보검색 서비스는 웹 상의 문서검색과 분류가 목적이기 때문에 획일적이고 정적인 특징을 가지고 있다(김원중, 조이기, 손철수, 2002). 또한 정보검색 서비스는 다수의 사용자를 대상으로 하기 때문에 각 기업의 요구를 개별적으로 만족시키지는 못한다. 대부분의 고객들은 자신의 요구사항이나 상품에 대한 추천 또는 불평사항을 인터넷에 산재한 수많은 사이트 중에서 선호하는 곳에 올리게 된다. 그러므로 기존의 정보검색 시스템을 사용하여 각 고객의 요구에 즉시 대응하기 위해서는 주기적으로 여러 사이트를 방문하여 모니터링하고 감지하여야 한다. 따라서 웹 상에 올라온 특정상품에 대한 내용이나 특정 사이트를 모니터링하여 꼭 필요한 정보를 제때에 파악할 수 있다면 매우 유용할 것이다.

이에 저자는 모니터링을 위한 기본 정보를 설정해 두면 자동으로 해당 정보를 감시하고 그 결과를 수집하여 실시간으로 사용자에게 전송하는 웹 모니터링 시스템을 개발하였다. 기존의 정보검색 시스템에서 데이터 수집을 목적으로 사용된 에이전트에 모니터링 설정 및 수행 기능을 부가하여 정보변화 감시(information change monitoring)와 정보 필터링(information filtering)을 수행할 수 있도록 하였다. 또한, 대화형으로 수행되는 에이전트를 개발하여 네트워크의 과부하를 억제할 수 있도록 하였으며, 모니터링 한 결과는 리포팅 시스템에 의해 사용자에게 보고서로 제공되며 전자메일로 통보할 수 있다.

본 논문의 구성은 다음과 같다. 먼저, II장에

서 웹 모니터링 관련 연구들을 분석한다. III장에서는 웹 모니터링 시스템의 전체 구성을 설명하고, 저자가 구현한 에이전트, 인텍싱, 리포팅 기능에 대하여 설명한다. IV장에서는 본 시스템을 인터넷 쇼핑몰에 적용하여 고객행동 및 응답데이터를 분석한 효용성 평가결과를 제시하며, V장에서 연구 결과를 정리하고 향후 연구 방향을 제시한다.

## II. 관련연구

‘모니터링’이란 무엇을 감시한다는 의미로 WWW(World Wide Web)이 발전하면서 필요성이 다방면에서 부각되고 있다. 현재 웹에 산재한 정보를 추출해 주는 에이전트에 관한 연구와 웹 페이지의 내용이 바뀌었을 때 사용자에게 자동으로 통보해 주는 웹 변경 탐지에 관한 연구들이 많지만 대부분 하나의 기능만을 달성하기 위한 연구들이고 여러 기능을 유기적으로 통합하여 제공하는 시스템은 드물다. 특히 이러한 웹 모니터링 시스템을 쇼핑몰 이용고객에 대한 응답소요시간 개선과 연관시켜 그 활용의 효과를 검증하고자 하는 연구는 미흡하다.

### 2.1 에이전트

에이전트는 대행자, 대리인 정도의 의미를 가지지만, 컴퓨터 분야에서는 작업을 대행해 주는 프로그램으로 해석될 수 있다. 학문적인 의미에서 에이전트는 분산 환경에서 작업을 수행하는 지적인 특성을 갖는 응용프로그램으로 정의되며, 연구 분야에 따라 다양한 의미로 해석

이 가능하다. 에이전트는 일반적으로 자율성 (autonomy), 사회성 (social ability), 반응성 (reactivity), 능동성 (proactivity), 시간연속성 (temporal continuity), 목표지향성 (goal-orientedness) 의 성질을 갖는다(도용태, 김일곤, 김종완, 박창현, 2001; 최무진, 황진열, 2005). 그리고, 에이전트는 응용분야에 따라 다양하게 분류되나 인터넷 상의 정보를 처리해 주는 에이전트는 크게 정보검색 에이전트, 정보필터링 에이전트, 정보통합 에이전트, 정보추출 에이전트의 네 가지로 분류할 수 있다. 정보검색 에이전트는 사용자가 원하는 정보를 찾아주는 역할을 수행하며 검색엔진이 대표적인 예가 된다. 정보필터링 에이전트는 사용자의 구미에 맞도록 정보를 가공하고 걸러준다. 정보통합 에이전트는 인터넷에서 제공되는 다수의 정보 사이트에서 사용자가 원하는 정보를 추출하여 하나의 형태로 제공하는 기능을 수행한다. 정보추출 에이전트는 한 문서에서 그 문서의 중심적 의미를 나타내는 구성요소를 인식하여 추출한다(최중민, 2000).

현재 에이전트는 활발하게 연구되어지고 있으며 다양하게 응용되고 있다. 그 중에서 로봇 에이전트는 전세계에 산재된 방대한 양의 정보를 스스로 추적하여 새로운 정보를 찾는다. 로봇 에이전트는 대부분 통계적인 목적이나 정보검색시스템을 위한 데이터 수집을 목적으로 사용되었다. 많은 정보를 수집하기 위해 더 높은 성능의 로봇 에이전트들이 제작되었고, 이러한 프로그램들이 팽창하면서 네트워크를 과부하 시키는 현상을 초래하게 되었다. 즉, 일반적인 사용자 과다로 인한 경우 이외에도, 로봇 에이전트에 의한 네트워크의 과부하가 문제점으

로 대두되었다(박철제, 조재균, 1999; 김경인, 김영진, 엄영익, 2004). 대부분의 에이전트 응용 시스템들은 보다 빠르게 정보를 추출하기 위해서 병렬적으로 수행시키는 방법을 사용한다. 즉 여러 개의 에이전트를 동시에 수행시키게 되고 이때 네트워크는 일관되게 높은 대역폭을 가지게 되어 같은 시간대의 다른 사용자는 느려진 네트워크 속도를 감수하게 된다. 이러한 문제를 해결하기 위하여 본 연구에서는 대화형으로 수행되는 에이전트를 구축하여, 무분별한 에이전트 사용을 제한함으로써 네트워크의 과부하를 억제하면서도 정보의 정확성과 신뢰성을 보장한다.

## 2.2 웹 변경 탐지

웹 변경 탐지는 인터넷의 활성화로 폭발적으로 증가하는 정보의 홍수 속에서 사용자들이 관심을 가지고 있는 정보의 변경을 곧바로 탐지하기 위한 연구이다. 웹 정보의 변화는 웹 문서의 내용이 변경되는 경우도 있고, 기존의 문서가 삭제되거나 추가되는 등 여러 가지가 있다. 기존의 웹 변경 탐지에 대한 연구를 특성과 역할에 따라 분류하면, 각각의 버전간의 변화를 감지하여 찾아내는 연구, 변화 감지된 정보를 사용자에게 알려주는 방식에 관한 연구, 시맨틱 웹을 이용하여 변화를 탐지하는 연구, 에이전트의 자동화된 서비스를 이용하여 변화를 탐지하는 연구 등 네 가지로 분류할 수 있다(조부연, 민영근, 이복주, 2006).

그리고 구현사례로서는, 교육용 포털 사이트의 사라진 사이트(dead site) 필터링을 위한 상호 협력적인 모니터링 에이전트의 구현(김영기,

한선관, 2000), 웹 문서를 의미 있는 단위로 구조화시키는 변경방법과 HTML 태그를 이용하여 분류시키는 방법을 제시하는 웹 사이트 콘텐츠 변경 모니터링 시스템(김원중, 조이기, 손철수, 2002), 웹에 대한 개별적인 모니터링을 처리하여 소요되는 자원 및 인력의 손실을 방지하는 웹 콘텐츠 변경 탐지 시스템(김영선, 장덕철, 2004), 시맨틱 웹과 온톨로지를 이용하여 의미에 변화가 있는 경우에만 웹 문서의 변경을 자동으로 사용자에게 알려주는 웹 변경 탐지 시스템(조부연, 민영근, 이복주, 2006) 등이 있다. 각 연구의 특징을 요약하면 <표 1>과 같다.

### Ⅲ. 시스템 설계 및 구현

본 시스템은 JAVA 기반의 환경에서 구현하였다. 웹 서버와 JSP 엔진을 결합하여, IIS는 HTML과 이미지 같은 스태틱 콘텐츠를 처리하고, Resin은 JSP와 Servlet 등의 JAVA 관련 콘텐츠를 처리하여 효율성을 높였다. 시스템은 수행될 작업에 따라 크게 에이전트, 인덱싱 및 리포팅 시스템으로 분산하여 구성하였다. 사용자가 모니터링 하고자 하는 사이트와 키워드를 설정하면 에이전트가 웹 문서, 이미지, 사운드 등을 모니터링하여 가져온다. 인덱싱 서버가 이를 분석해서 그 결과를 저장하고, 리포팅 시스템에 의해 보고서가 생성된다. 생성된 보고서는

<표 1> 각 연구의 특징

구 분	특 징
협력적인 모니터링 에이전트 구현 (김영기, 한선관, 2000)	<ul style="list-style-type: none"> <li>· 교육용 포털 사이트를 대상으로 이동 에이전트를 이용하여 사라진 사이트 모니터링</li> <li>· 에이전트를 이용함으로써 수행속도 향상, 모니터링 서버의 최대 활용 및 상호협력 작업을 통한 효율성 증가</li> </ul>
콘텐츠 변경 모니터링 시스템 구현 (김원중, 조이기, 손철수, 2002)	<ul style="list-style-type: none"> <li>· HTML 문서의 의미적 변경을 단순 변경, 집합적 변경 및 갱신 변경으로 모니터링하고 그 결과를 사용자에게 통지</li> <li>· 웹 문서를 의미있는 단위로 구조화시키는 변경 방법과 HTML 태그를 이용하여 의미있는 단위로 분류시키는 방법 제시</li> </ul>
콘텐츠 변경 탐지 시스템 구현 (김영선, 장덕철, 2004)	<ul style="list-style-type: none"> <li>· 콘텐츠에 대한 불법적인 침입을 차단할 수 있는 모니터링 기능 수행을 통해, 콘텐츠 보안 취약성과 정보노출을 보호</li> </ul>
웹 변경 탐지 시스템 구현 (조부연, 민영근, 이복주, 2006)	<ul style="list-style-type: none"> <li>· 시맨틱 웹과 온톨로지를 이용하여 의미중심의 변경 탐지</li> <li>· 빠른 탐지 및 성능향상을 위하여 서버와 클라이언트의 역할 분할</li> </ul>
본 연구	<ul style="list-style-type: none"> <li>· 에이전트에 모니터링 설정 및 수행 기능을 부가하여 정보변화 감시와 정보필터링 수행</li> <li>· 대화형으로 수행되는 에이전트를 구축하여 무분별한 에이전트 사용을 제한함으로써 네트워크의 과부하 억제</li> <li>· 인터넷 쇼핑몰에 적용하여 이용고객에 대한 응답소요시간 평가</li> </ul>

<표 2> 시스템 세부기능

구 분	세부 기능
에이전트	<ul style="list-style-type: none"> <li>· 기본정보 및 고급정보 설정</li> <li>· 설정된 정보를 바탕으로 웹 문서를 수집하여 인텍싱 시스템에 제공</li> </ul>
인텍싱	<ul style="list-style-type: none"> <li>· 수집된 문서에서 불용어 제거</li> <li>· 사전을 이용하여 색인어를 추출하고 색인어의 출현빈도 분석</li> </ul>
리포팅	<ul style="list-style-type: none"> <li>· 사용자가 제공한 키워드를 기반으로 데이터를 분석하여 보고서 생성</li> <li>· 보고서 전송</li> </ul>

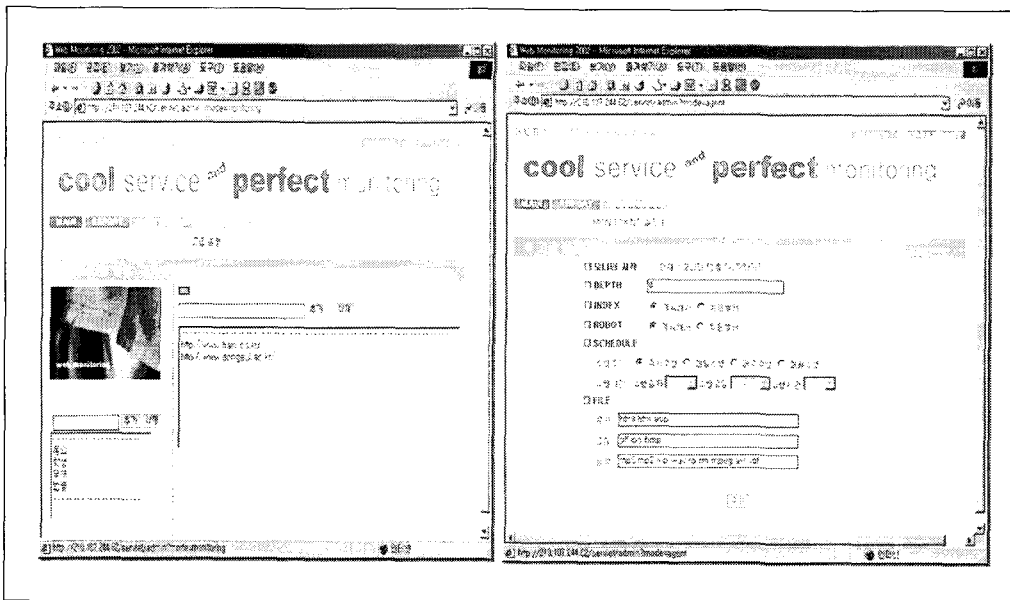
전송 기능을 통해 사용자에게 실시간으로 제공된다. 구성별 세부 기능을 요약하면 <표 2>와 같다.

### 3.1 에이전트

#### 3.1.1 모니터링 설정

모니터링 기능을 수행하기 위해서는 대상 사이트와 키워드를 지정하는 기본정보 및 에이전

트의 구동을 위해 필요한 Depth, 수행주기, 검색대상 파일 등의 고급정보를 설정한다. 에이전트는 설정된 정보를 바탕으로 모니터링을 수행하고 그 결과를 인텍싱 시스템에 넘겨준다. 본 시스템에서는 모니터링 설정 기능을 통하여 무분별한 에이전트의 사용을 제한함으로써 네트워크의 과부하를 억제하였다. 기본정보 및 고급정보 설정을 위한 구현 화면을 <그림 1>에 나타내며, 각 설정 요소들의 세부기능은 <표 3>



<그림 1> 모니터링 설정 화면

<표 3> 모니터링 설정 세부기능

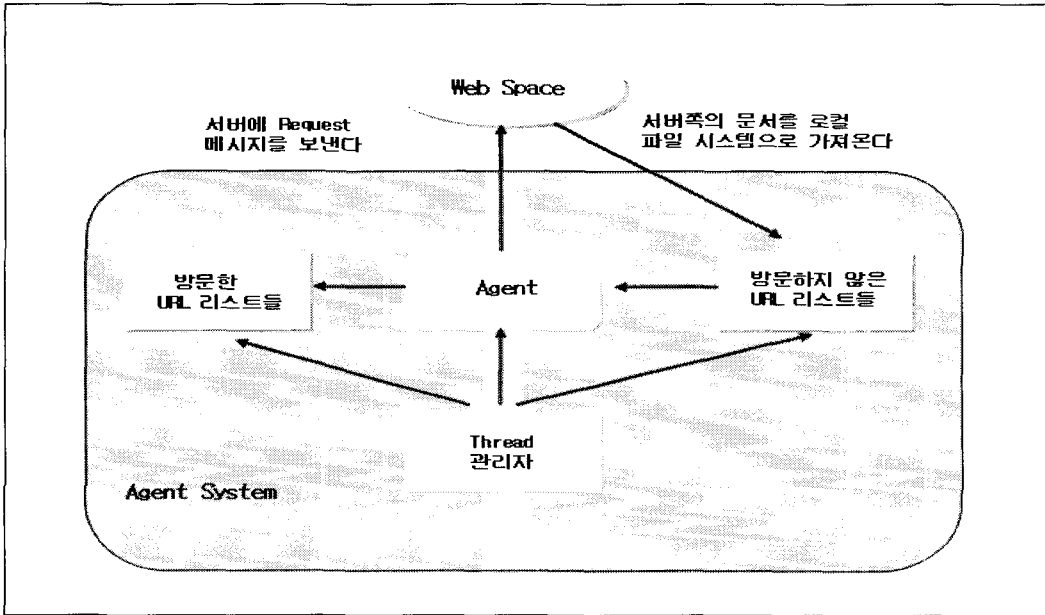
구 분		세부 기능
기본 정보	키워드	<ul style="list-style-type: none"> <li>· 모니터링 대상 키워드의 추가·삭제</li> <li>· 설정된 키워드에 한하여 모니터링을 수행</li> <li>· 예 : 키워드로 “박지성”을 입력하면 웹 문서 내용 중 “박지성”이 있는 웹 문서를 모니터링</li> </ul>
	URL	<ul style="list-style-type: none"> <li>· 모니터링 대상 사이트의 URL 추가·삭제</li> <li>· 설정된 URL에 한하여 모니터링을 수행</li> <li>· 예 : “http://www.deu.ac.kr”을 입력하면 “동의대학교”의 웹 사이트를 모니터링</li> </ul>
고급 정보	모니터링 시작/멈춤	<ul style="list-style-type: none"> <li>· 에이전트의 구동 및 멈춤 기능</li> </ul>
	Depth	<ul style="list-style-type: none"> <li>· 모니터링 대상 URL의 하위 디렉토리 깊이를 지정</li> <li>· 예 : “3”을 입력하면 모니터링 할 URL의 3개 하위 디렉토리 수준까지 모니터링</li> </ul>
	Index	<ul style="list-style-type: none"> <li>· 기존 인덱싱 작업에 이어서 인덱싱을 수행하려면 “계속해서”를 선택하고, 기존의 인덱싱 작업을 삭제하고 처음부터 시작하려면 “처음부터”를 선택</li> </ul>
	Robot	<ul style="list-style-type: none"> <li>· 모니터링 수행 시, 로봇 에이전트 배제 규칙의 적용 여부</li> <li>· 로봇 에이전트가 구동 될 때 기존에 방문한 URL 이후부터 모니터링을 수행 할 것인지, 처음의 URL 부터 모니터링을 수행 할 것인지를 선택</li> </ul>
	Schedule	<ul style="list-style-type: none"> <li>· 모니터링 수행주기 및 시간 설정</li> <li>· 즉시수행, 매일수행, 매주수행, 매월수행</li> <li>· 수행일, 수행요일, 수행시간의 설정</li> </ul>
	File	<ul style="list-style-type: none"> <li>· 모니터링 대상 파일의 확장자를 지정</li> <li>· 문서, 그림, 음악의 각각에 해당하는 파일 확장자를 입력</li> </ul>

과 같다.

### 3.1.2 모니터링 수행

모니터링은 에이전트에 의해 수행되며, 수집된 문서는 로컬 파일 시스템에 일정한 형식에 따르는 파일명으로 저장된다. 쓰레드 관리자에 의해 수행되는 에이전트는 설정된 고급정보에 따라 방문한 URL 이후부터 모니터링을 수행할 것인지, 처음의 URL부터 모니터링을 수행할 것인지를 결정한다. 쓰레드 관리자가 에이전트에게 수행해야 될 URL을 넘겨주면, 에이전트

는 자체적으로 URL을 파싱하여 호스트, 포트 번호, 레퍼런스를 추출해 낸다. 이때 표준 URL 형식에 대한 타입으로 변환하여 소켓을 통해 서버에 메시지를 전달한다. 에이전트를 크게 소켓 연결부, Request 수행부, Response 수행부, URL 파싱, URL 추출, 인덱스 관리 등으로 구분하였다. 소켓 연결부는 자바의 java.net.Socket 클래스를 이용하여 구현하였다. Request/Response 수행부는 HTTP 규약에 따른 포맷으로 구성되며, URL 파싱은 java.net.URL 클래스로 구현하였다. 에이전트 수행 알고리즘의 기본 원리는



<그림 2> 에이전트의 동작방식

소켓을 이용하여 웹 서버에 접속한 뒤에 스트림을 통하여 Request/Response 규약에 따른 통신을 이용한다. 일단 연결이 설정되면 스트림으로 Request 메시지를 보내고, 메시지 전달이 끝났음을 알리면, 서버는 다시 스트림을 통해 Response 메시지를 되돌려 준다. 하지만 여기서 상태 코드를 분석하여 유용한 정보일 때 만 파일로 저장한다. <그림 2>는 에이전트의 동작방식을 보여준다.

### 3.2 인덱싱

인덱싱 시스템은 에이전트가 수집한 문서를 분석하여 색인어를 추출하고 한 문서 안에서 색인어의 출현 빈도를 분석한다. 인덱싱 시스템에서 추출하는 색인어를 사용하여 모니터링 결과를 제공하게 됨으로 정확한 색인어 추출은 모니터링 시스템의 성능에 중요한 역할을 한다.

특히, 한글문서는 구조적인 특성상 띄어쓰기의 오류 및 조사 활용에 대한 변화가 다양하기 때문에 의미 있는 단어를 찾아내기 위해서 형태소 분석은 필수적인 요소라 할 수 있다.

#### 3.2.1 특징

인덱싱 시스템의 특징을 요약하면 다음과 같다.

- (1) 전체 시스템에서 가장 큰 부하를 가져오는 부분이다.
- (2) 본 인덱싱 시스템은 한글과 영문에 대해서 독립적으로 처리한다. 한글의 경우 복잡한 수행과정을 거친다. 즉 HTML 태그 제거 → 형태소 분석 → 명사사전 검색 → 스크립트 생성을 통해서 의미 있는 명사를 찾는다. 영문의 경우 지식 베이스를 이용해서 인덱싱을 한다. 즉 새로 생긴 용어들(ex: hotjava, intranet, netscape,

altavista ...)에 대해서 능동적으로 반응을 한다. 최초 새로운 토큰에 대해서 지식 베이스에 등록을 한 뒤 다른 문서에서 3번 이상 같은 토큰이 발견될 경우 새로운 용어로 등록시킨다. 이런 경우 메일 주소나 호스트명 등도 3번 이상 발견되면 새로운 용어로 간주하고 실제 인덱스 스크립트를 만들 때 참조하게 된다. 현재 영문에 대해 새로운 용어를 포함하여 200,000단어 이상으로 검색이 가능하다. 한글의 경우 사전에 등록된 50,000단어를 사용한다. 한글의 경우 복잡한 구조적 특성으로 이런 지식 베이스를 유지하기에 많은 무리가 따른다. 결국 한글 처리 능력은 사전에 종속된다고 해도 과언이 아니다.

- (3) 인덱스의 속도를 향상하기 위해서 조사사전과 명사사전에 대해서 색인 파일을 유지한다. 한글의 경우 초성을 분리해 낸 후, 분리된 초성에 따라서 색인파일을 조사하여 실제 사전파일에 대한 검색 속도를 향상시킨다. 조사의 경우 초성으로 분리해 내는 것은 의미가 없다. 한글은 조사와 결합하기 위하여 한 글자 이상의 어간을 가지게 되고, 결국 토큰의 길이에서 조사 길이의 한계를 찾아낼 수 있다. 이런 조사 길이에 대한 정보를 색인으로 유지하고 토큰에 대해서 조사색인을 참조하여 4,000개의 조사사전과 비교하여 불용어를 제거한다.

### 3.2.2 형태소 분석

형태소 분석은 한국어 어절을 형태소라는 의

미적 최소 단위로 분리하는 작업을 말한다. 형태소 분석 과정은 입력 문서에서 한 단어씩 추출하여 입력된 단어의 한글 코드를 내부 코드인 조합형 한글 코드로 변환하는 전처리 단계를 거치며, 형태소 분석 결과는 시스템의 요구 사항에 맞게 출력 결과를 조절하는 후처리 단계의 과정으로 되어 있다.

본 연구에서는 HTML 문서에 대한 태그 파싱 처리가 이루어진 결과에 대해서 형태소 분석을 수행하여 유용한 단어를 찾아내는 방법을 사용하였다. 우선 불용어에 대한 선 처리가 이루어져야 한다. 불용어는 토큰 단위로 구분된 문자열에 대해서 의미 없는 조사들에 해당하며, 이때 조사사전을 이용한 매칭 기법을 사용한다. 조사에 대한 검사가 잦고 조사의 개수가 많지 않은 만큼 조사사전은 메모리에 상주시켜 사용함으로써 매칭 수행속도를 향상시켰다. 하나의 토큰에 대해서 조사를 찾는 방법은 토큰의 뒤에서 부터 매칭을 시도하여 제거하는 방법을 사용한다. 이때 속도를 높이기 위해서 인덱스를 유지하는데 한글 초성에 대한 인덱스는 효과를 보기가 어렵다. 본 연구에서 사용된 방법은 음절의 수를 이용한 인덱스 기법으로서 토큰을 대상으로 조사사전과 비교하는 방법을 채택하고 있다.

불용어 제거가 끝난 결과에 대해서 다시 명사사전과 비교해서 의미 있는 단어를 추출하는 작업이 필요하다. 명사사전은 방대한 양의 데이터를 유지하기 때문에 메모리에 상주할 수 없다. 그래서 외부 파일과 더불어 인덱스를 유지하여야만 한다. 인덱스의 구성은 한글 초성에 바탕을 두고 있으며, 불용어가 제거된 토큰에 대해서 첫 문자의 초성을 구한 뒤, 각 초성에



해당하는 위치에서 부터 매칭 작업에 들어가게 된다. 이 부분이 형태소 분석의 전체 과정에서 가장 많은 부하를 가져오는 부분으로서 인덱스 구조를 사용하여 수행속도를 증가 시켰다.

형태소 분석 과정에서 추출된 키워드를 대상으로 질의를 통한 검색을 지원하기 위하여 스크립트를 생성해야 한다. 스크립트에 포함되는 정보는 제목, 본문요약, URL, 웹 문서 크기, 이미지 및 사운드의 URL 정보 등이다.

### 3.3 리포팅

리포팅은 보고서를 생성하기 위한 기능이다. 에이전트가 위의 과정을 거쳐 모니터링 한 결과를 바탕으로 사용자가 현재의 상황을 가장 잘 파악할 수 있도록 간단 명료하고 보기에 좋도록 보고서를 생성하여야 한다.

사용자에게 제공되는 보고서는 모니터링 대상 키워드에 따라 그 결과를 달리 제공하는 것을 원칙으로 하며, 모니터링 결과를 한 눈에 파악할 수 있는 그래프, 요약된 모니터링 결과 리스트, 모니터링 결과 원문 정보를 포함하도록 개발하였다. 리포팅은 크게 데이터 분석기능, 보고서 생성기능 및 전송기능으로 분류할 수 있다.

에이전트로 검색한 데이터를 인덱싱 시스템을 통해 스크립트를 생성하면 일차적인 모니터링 결과가 생성된 것이다. 데이터 분석기능은 이러한 일차적인 모니터링 결과를 분석하여 가공하는 처리과정을 다룬다. 정보 수집 단계에서 생성된 스크립트 파일을 사용자가 지정한 키워드를 기반으로 하여 분석한다.

보고서 생성기능은 앞에서 분석된 데이터를 기반으로 이루어지는데, 사용자가 설정한 키워

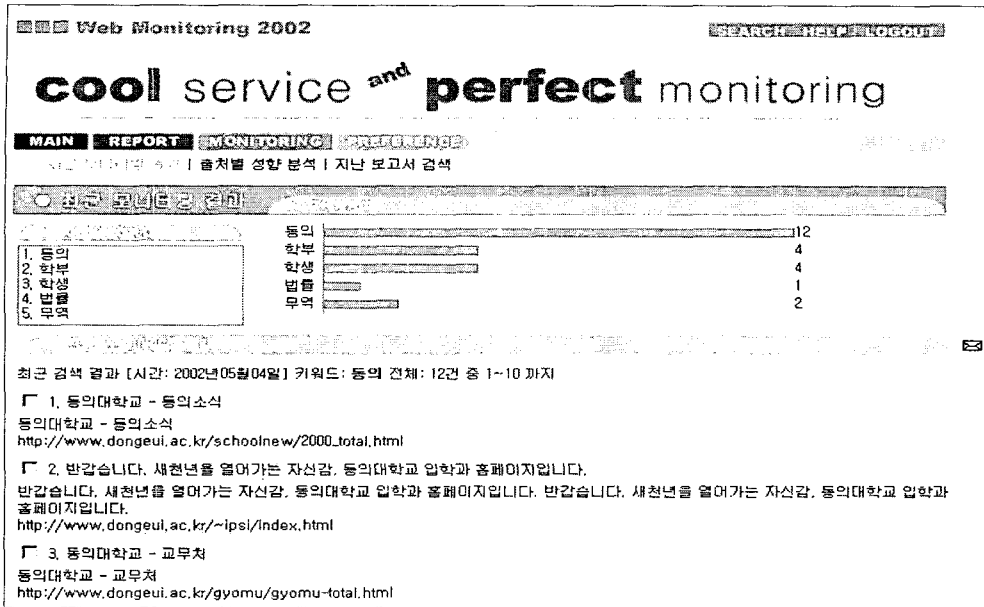
드로 질의가 수행되어 요약 보고서를 생성한다. 보고서는 모니터링 일자별로 생성되며, 파일 시스템에 저장된다. 따라서, 차후 일자별 보고서 검색이 가능하다. 보고서는 그래프 정보, 모니터링 결과 수치 정보, 모니터링 결과 정보 등을 포함한다.

전송기능은 모니터링 된 결과 보고서를 사용자에게 자동적으로 제공해 주는 기능으로, 웹 모니터링 시스템을 통한 직접적인 Push 기능과 E-mail 전송 기능으로 나뉘어진다. 웹 모니터링 시스템에 사용자가 로그인하는 경우 업데이트 된 모니터링 결과 보고서와 원문 리스트를 보여 준다. 사용자는 필요에 따라 E-mail로 보고서를 전송할 수 있는데, 웹 모니터링 시스템에서는 자바 웹 메일을 이용하여 보고서가 E-mail로 전송되어질 수 있도록 시스템을 구현하였다.

## IV. 효과 분석 및 평가

본 연구에서 설계한 웹 모니터링 시스템을 Window 2000 Professional 환경에서 JDK(Java Development Kit)를 사용하여 구현하였다. <그림 3>은 모니터링 설정 화면에서 “동의”, “학부”, “학생”, “법률”, “무역”의 5개 키워드를 입력하고, 대상 URL을 “동의대학교” 웹 사이트로 설정하여 모니터링 한 결과 화면을 나타낸다.

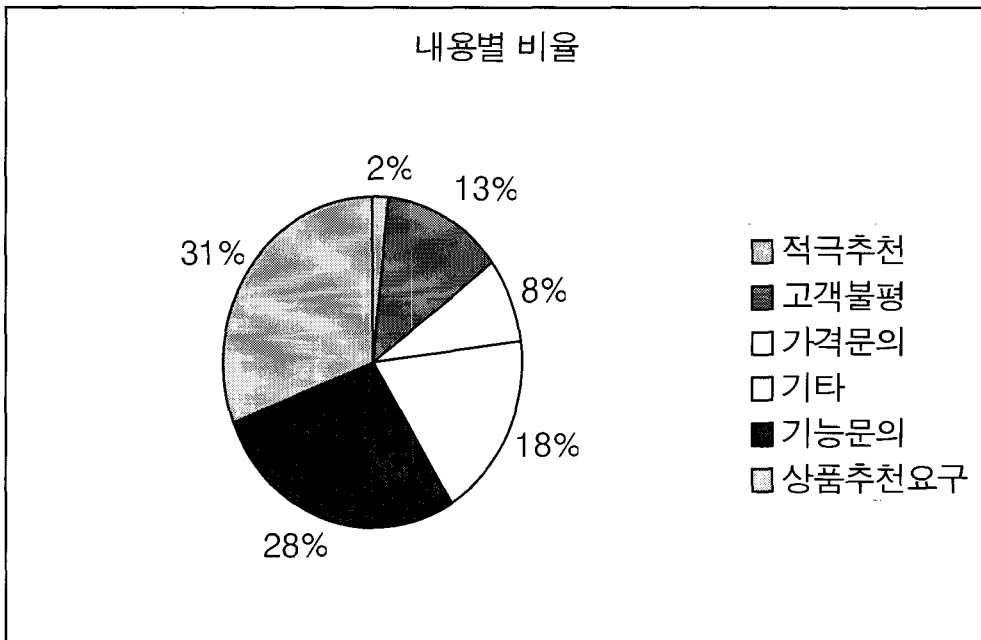
본 시스템은 기존 정보검색 시스템의 한계점을 극복하기 위해, 필요한 정보만을 즉시에 수집하여 여러가지 형태로 제공할 수 있는 기능을 갖추어서 다방면에 활용할 수 있도록 하였다. 웹 상의 근거 없는 루머 확산, 개인 이미지 손상과 관련된 글들로 인하여 여론에 민감



<그림 3> 모니터링 결과 화면

한 정치인이나 연예인 뿐만 아니라 기업들의 제품 이미지 관리에도 신속하고 적극적으로 대

응 가능하다. 실제 각 기업에서는 고객서비스 향상 차원에서 고객센터 운영에 많은 비용을



<그림 4> 내용별 비율

지불하고 있으나, 웹 상의 고객응대 측면에서는 만족할 만한 성과를 거두지 못하고 있다. 특히, 쇼핑몰 업체에서는 경쟁이 치열해짐으로 인하여 고객의 불평행동이나 질의에 대한 응답소요시간의 개선이 기업의 이윤증가에 막대한 영향을 미치고 있다.

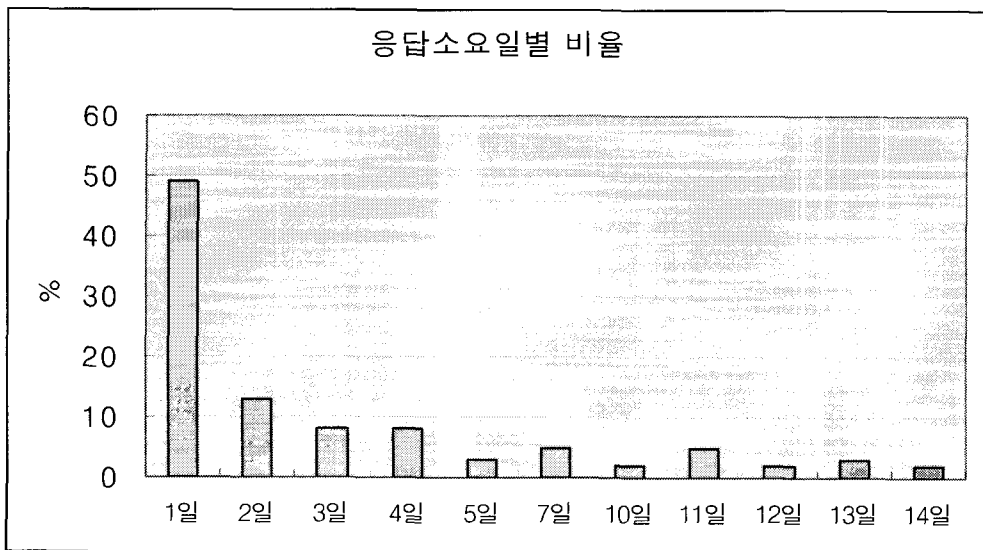
따라서, 본 시스템을 이용하여 쇼핑몰 이용 고객에 대한 응답소요시간을 평가하고 그 활용 효과를 검증하기 위하여, “N사이트”, “S사의 LCD 모니터”를 기본정보로 설정하여 웹 모니터링 시스템을 가동하였다. 그 결과, 2004년 9월부터 2006년 8월까지 793건의 웹 문서를 수집하였다. 하나의 글에 대한 평균 조회수를 분석해 보면 573번으로 웹 상에서 고객들의 행동이 얼마나 큰 파급 효과를 가지는지를 알 수 있다.

내용을 분석해보면 LCD 모니터 중 어느 것이 적당한지 추천을 요구하는 글들이 전체의 31%로 가장 많았으며, 제품의 기능에 대한 질

문은 28%로 두 번째를 차지하고 있다. 고객이 구매한 물품을 적극적으로 추천한 경우는 2%로 상대적으로 적은 양을 차지하고 있다. 이는 고객들이 구매한 물품에 대하여 고객만족을 확실하게 느끼지 않는 한 적극적으로 행동하지 않는 것으로 판단된다. <그림 4>는 내용별 비율을 나타낸다.

고객만족과 고객에 대한 응답소요시간은 깊은 상관관계를 가지고 있으며 고객만족을 달성하기 위해서는 웹 상의 글들에 즉시 응답 할 필요가 있다. 그러나 현재의 환경에서는 평균 응답 소요일이 3.4일이 걸린 것으로 나타나고 있다. <그림 5>는 응답소요일별 비율을 나타낸다.

전체적으로 하루 이내에 응답한 비율이 49%로 가장 많으나, 51%의 고객들에 대해서는 응답에 이틀 이상이 소요된 것으로 나타나고 있다. 특히 10일 이상 소요된 데이터도 14%로 이는 고객들의 이탈 현상으로 나타날 것으로 판단된다. 이는 현재의 상황이 가지고 있는 한계



<그림 5> 응답소요일별 비율

로 이를 극복하기 위해서는 본 시스템의 모니터링 설정 기능을 이용하여 특정 상품에 대한 내용이나 특정 사이트를 모니터링하여 필요한 정보를 제때에 파악하여 대응함으로써 개선될 수 있다.

## V. 결론

본 연구에서는 웹 상의 정보를 효율적으로 모니터링하고 그 결과로 보고서를 생성하여 실시간으로 사용자에게 전송하는 시스템을 개발하였다. 시스템 개발의 측면에서 본 연구는 에이전트에 모니터링 설정 및 수행 기능을 추가하였다. 고급정보 설정을 통하여 에이전트의 구동을 제어할 수 있도록 하였으며, Index 및 Robot 정보설정 기능을 이용하여 무분별한 에이전트 사용을 제한함으로써 네트워크의 과부하를 억제하였다. 시스템의 효율성 측면에서 본 연구는 쇼핑몰 이용 고객들의 요구사항이나 상품에 대한 추천 또는 불평사항에 즉시 대응하지 못하는 한계점을 파악하였다. 이는 본 시스템을 이용하여 필요한 정보를 모니터링 함으로써 개선될 수 있다.

본 시스템은 쇼핑몰 분야 이외에도, 인터넷 상에서 발생할 수 있는 개인과 기업의 이미지 실추, 명예훼손, 디지털 콘텐츠의 무단 복제 및 배포 등으로 발생할 수 있는 피해를 미리 예방하고, 이미 문제가 발생하였을 경우 빠른 시간 내에 효율적으로 해결하여 피해를 최소화 할 수 있다. 인터넷의 부정적인 측면으로 부각되고 있는 유해정보의 범람, 기업과 관련된 악성루머 등은 상품의 판매, 기업의 이미지 및 가치에 막

대한 영향을 미치고 있다. 그러므로 웹 모니터링 시스템을 통해 기업은 고객만족을 달성할 수 있으며 이로 인하여 경쟁기업보다 우위를 점할 수 있게 될 것이다.

본 연구의 한계점으로는, 쇼핑몰 이용고객에 대한 응답소요시간 개선의 효과를 광범위하게 측정할 필요성이 있으나 이를 실현하지 못한 점을 들 수 있다. 상업적인 쇼핑몰에 저자가 개발한 웹 모니터링 시스템을 결합하여 운영함으로써 시스템의 효과를 검증 할 수 있는 다양한 데이터 수집이 가능 할 것이다. 그러나 이는 쇼핑몰 업체와의 전략적 제휴 등을 고려하여 장기적으로 실현해야 할 사항으로 판단된다. 또한, 마케팅 분야에서 많은 연구가 이루어지고 있는 고객충성도 향상에도 본 시스템이 효과를 미칠 수 있을 것으로 판단된다. 고객의 불평행동 개선을 통한 고객만족, 신뢰, 몰입 등이 고객의 충성 행위에 큰 영향을 미치고 있다. 그러나 고객충성도와 본 연구를 연계시키기 위해서는 광범위 한 후속 연구가 필요하므로 이는 향후 연구를 통하여 그 효과를 검증 할 수 있을 것이다.

## 참고문헌

- 김경인, 김영진, 엄영익, “순수 P2P 네트워크 환경에서 에이전트 이주를 위한 자원 예약 기반 동적 부하 균형 기법,” 정보처리학회논문지A, 제11-4권, 제4호, 2004, pp. 257-266.
- 김성학, 이창훈, “웹 로그 분석을 이용한 추천 에이전트의 개발,” 정보과학회논문지: 기술교육, 제2권, 제1호, 2005, pp. 60-66.

김영기, 한선관, “이동 에이전트를 이용한 협력적인 모니터링 에이전트의 설계 및 구현,” 한국정보교육학회논문지, 제4권, 제1호, 2000, pp. 24-31.

김영선, 장덕철, “웹 콘텐츠 변경 탐지 시스템의 설계 및 구현,” 멀티미디어학회논문지, 제7권, 제1호, 2004, pp. 91-97.

김용기, 양재영, 최중민, “사용자 인터페이스 에이전트를 통한 정보추출 규칙의 자동 생성,” 정보과학회논문지: 소프트웨어 및 응용, 제31권, 제4호, 2004, pp. 447-456.

김원중, 조이기, 손철수, “웹 사이트 콘텐츠 변경 모니터링 시스템,” 한국해양정보통신학회논문지, 제6권, 제4호, 2002, pp. 505-512.

김태훈, 최중민, “사용자 편의의 인터넷 정보검색을 위한 지능형 웹 브라우징 에이전트,” 정보과학회논문지(B), 제25권, 제7호, 1998, pp. 1064-1078.

도용태, 김일곤, 김종완, 박창현, 인공지능 개념 및 응용, (주)사이텍미디어, 2001.

박준철, “인터넷 쇼핑몰 이용자의 고객만족이 신뢰, 몰입, 고객충성행위에 미치는 영향,” 경영정보학연구, 제13권, 제3호, 2003, pp. 131-149.

박철제, 조재균, “실시간 인덱싱을 위한 대화형 로봇 에이전트의 설계 및 구현,” 동의논집: 인문사회과학편, 제30집, 1999, pp. 745-760.

이승익, 조성배, “웹기반 대화형 에이전트,” 정보과학회논문지: 컴퓨팅의 실제, 제9권, 제5호, 2003, pp. 530-540.

조부현, 민영근, 이복주, “시맨틱 웹을 이용한

웹 변경 탐지 시스템,” 정보처리학회논문지B, 제13-B권, 제1호, 2006, pp. 21-26.

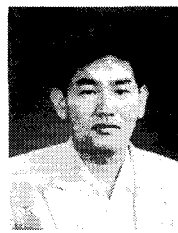
최무진, 황진열, “비교쇼핑 에이전트의 설계와 구현,” *Information Systems Review*, Vol. 7, No. 1, 2005, pp. 97-113.

최중민, “인터넷 정보 추출 에이전트,” 정보과학회지, 제18권, 제5호, 2000, pp. 48-53.

현정석, 하환호, “인터넷 쇼핑몰에 대한 이용고객의 충성도와 불평행동과의 관계,” 대한경영학회지, 제19권, 제3호, 2006, pp. 781-807.

Bettencourt, L. A., “Customer Voluntary Performance: Customers As Partners in Service Delivery,” *Journal of Retailing*, Vol. 73, No. 3, 1997, pp. 383-406.  
<http://www.cywatcher.com>.

**박철제(Chul-Jae Park)**



연세대학교를 졸업하고, 일본 와세다대학교에서 정보과학전공 석사 및 박사학위를 취득하였으며, 현재 동의대학교 경영정보학과 부교수로 재직하고 있다. 와세다대학교 미디어네트워크센터 특별연구원, 현대정보기술연구소 책임연구원 및 미국 University of California Irvine에서 연구교수로 근무하였으며, 관심분야는 Web service, e-Learning, 에이전트 등이다.

**조재균(Jae-Gyeun Cho)**

연세대학교 응용통계학과를 졸업하고, 한국과학기술원에서 산업공학전공 석사 및 박사학위를 취득



하였으며, 현재 동의대학교 e  
비즈니스학과 부교수로 재직  
하고 있다. 한국전자통신연  
구원 선임연구원으로 근무하  
였으며, 관심분야는 프로젝  
트 관리, 인터넷비즈니스 모  
델 및 전략 등이다.

<Abstract>

## A Study on the Implementation and Practical Application of Web Monitoring System

Chul-Jae Park · Jae-Gyeun Cho

The web monitoring is a procedure to detect and monitor the change of a particular information from the web site. In this paper, we propose a monitoring system to survey the response time for the customers on an internet shopping mall. This system we designed, decreases the network overload and guarantees the accuracy as well as the reliability of the information by using an interactive agent. This interactive agent we constructed, is written in Java and controlled by the high-level information such as the index or the robot. From the N web site, we extracted the web monitoring results of keyword "LCD Monitor" and verified our system's efficiency. The performance of the system is tested and the result is illustrated.

**Keywords:** Web Monitoring, Agent, Indexing

\* 이 논문은 2006년 11월 29일 접수하여 3차 수정을 거쳐 2007년 3월 20일 게재 확정되었습니다.