韓國컴퓨터情報學會 論文誌
第15卷 第6號, 2010. 6.

2010-15-6-5-5

검색엔진 최적화를 위한 웹사이트 가시성 분석에 관한 연구

윤 선 희*

Study on Analysis of Website Visibiliy for Search Engine Optimization

Sunhee Yoon*

요 약

인터넷이 주요 마케팅과 영업의 통로로 사용됨에 따라 검색 엔진 결과의 상위 순위가 웹 사이트들 사이의 핵심 경쟁대상이 되어가고 있다. 검색엔진에서 웹사이트의 상위 순위를 유지하기 위해 다양한 방법이 존재하며 일반적으로 유기적인 코딩이나 검색엔진 최적화에 많은 비용을 투자한다. 본 논문의 목적은 웹사이트 가시성을 충족하지 못할 경우 검색엔진색인에서 순위가 뒤떨어지거나 완전히 제거될 수 있기 때문에 검색엔진 최적화(Serch Engine Optimization : SEO)를 위해 웹사이트 가시성을 고려하여 웹페이지를 설계할 때 부정적 요소로써 제거되어져야 할 요인들을 인식하여 순위를 제시하고자 한다. 본 논문에서 제안한 웹사이트 가시성의 부정적 요인을 인식하고 순위를 결정한 실험은 기존의 웹사이트 가시성 분석 모델을 기반으로 이론과 실험을 통해 제공되어졌다. 본 논문에서 분석된 모델들은 각 모델의 방법론을 기반으로 점수로 표현되거나 계량화되었으며 실험을 통해 부정적 요인으로 10개의 항목이 선정되어 높은 점수로 순위가 결정되었다. 따라서 웹사이트를 설계할 때 부정적인 요인들인 높은 순위의 항목들을 배제하여 설계함에 따라 웹사이트가 검색엔진 색인 대상에서 제거되지 않도록 고려되어져야 한다.

Abstract

The Internet has become a major channel of business marketing and sales, and there is a core competitive object between websites for a high position ranking in search engine results. There are various ways to maintain the high position ranking of website involving the development of componental coding or the expensive investment for the search engine optimization. The purpose of this paper is proposed to identify and rank the negative elements of website visibility to get rid of those elements when website designer designs the webpage. Website can be removed from indices of search engines when they are not satisfied for search engine optimization. The proposed experiments that are identified and ranked the negative elements of website visibility in this paper are based on the theories and experiments of existing website visibility models. The experimental analyses in this paper are scored and normalized based on methodologies of those models and 10 highest negative elements are ranked through the analyses. Therefore when website is designed, these highest negative elements should be avoided so website can not be removed in the indices of search engines.

▸ Keyword : 검색엔진(Search Engine) 최적화(Optimization), 웹사이트(Website), 가시성(Visibility)

• 제1저자 : 윤선희
• 투고일 : 2010. 04. 20, 심사일 : 2010. 04. 26, 게재확정일 : 2010. 05. 12.
* 숭의여자대학 인터넷정보과 교수
※ 본 논문은 2009년도 숭의여자대학 교내학술연구비 지원에 의해 연구되었음.

Ⅰ. 서 론

현대 사회에서 인터넷은 기업이나 우리의 일상생활에서 빠질 수 없는 중요한 영향을 미친다. 인터넷에서 가장 많이 사용되는 분야는 이메일과 검색으로 인터넷 사용자의 80% 이상이 검색에서 발생하는 트래픽이라는 주장이 제기되고 있다[1]. 또한 온라인 통신과 관련된 정보들을 찾기 위한 검색엔진의 사용은 우리가 인터넷을 사용하는 방법들에서 현저히 나타난다. 분석가들에 의하면 검색자들의 평균 90%가 세 번째 검색엔진결과 페이지(Search Engine Result Page:SERP)를 참고하지 않는다는 결과가 나왔다[1]. 온라인 구매 결정에 있어서 검색과정과 검색결과의 사용이 중요한 역할을 함에 따라 웹사이트를 구축할 때 SERP의 상위를 확보하기 위한 웹사이트를 만드는 것이 무엇보다 중요하다. 웹사이트 설계자들은 웹페이지를 설계함에 있어서 유기적인 디자인 방법을 선택하거나 검색엔진 결과페이지의 상위 순위를 유지하기 위한 비용 지불을 감수해야 한다. 현재까지 검색엔진 검색결과의 상위 순위를 유지하기 위해 반드시 필요한 요인이나 제거되어야 할 요인의 지침이 제공되고 있으나 학문적 연구 모델을 기반으로 실험을 통한 순위 시스템은 만들어 지지 않고 있다. 본 논문의 목적은 검색엔진에서 제거될 수 있는 요인들을 웹 가시성 모델의 연구 방법론을 기반으로 실험을 통해 순위를 생성하는 것이다. 또한 실험을 통해 생성된 순위가 높은 부정적 요인들은 웹사이트를 설계할 때 우선적으로 제거되어져야 할 항목으로 주목할 수 있다.

Ⅱ. 관련 연구 분석

2.1 배경

검색엔진 최적화(SEO)는 허수의 방문자에 의한 상위 순위나 방문자수의 증가를 조율하는 처리과정으로 기업의 이익을 극대화하기 위한 온라인 마케팅의 핵심이며 기초 작업으로써 SEO 성과를 진단하고 평가하기 위한 다양한 방법이 존재한다. 그 중 로그 분석을 통해 검색엔진 최적화를 진단 평가한 예를 사용해 SEO가 기업 영업 이익에 미치는 영향력을 분석해 본 결과 SEO 적용 전에는 50% 이상이 영업이익에 불필요한 트래픽이 존재하였으나 SEO를 세밀히 적용함에 따라 영업이익에 직접 영향을 미칠 수 있는 네이버와 같은 의미 있는 검색엔진의 방문율이 높아짐을 알 수 있다<그림 1> [2][3][4].

소스	SEO 적용 전	SEO 기초작업 적용 1주일 후	SEO심층작업 적용 1달 후
직접트래픽	54%	4.64%	8.02%
네이버	18%	31.25%	43.4%
야후	10%	7.06%	6.37%
다음	8%	11.29%	15.09%
구글	6%	40.52%	24.29%

그림 1. 검색엔진최적화 진행단계 비교 테이블
Fig. 1. Comparison for SEO Process

기업의 입장에서는 인터넷 이용자에게 인기가 있는 검색엔진의 방문율이 향상되는 것은 구매율의 증가나 영업 이익률의 증가를 의미함에 따라 웹사이트 가시성에 검색엔진 결과의 상위 순위를 유지하기 위한 방법론에 집중적인 관심을 나타낸다. 따라서 웹사이트가 실제 컨텐츠에 따라 검색 순위가 결정되는 것이 아니라 스팸이나 부정적인 항목들에 의해 결정되지 않도록 웹사이트 가시성 요인을 정확히 분석하여 검색엔진 최적화가 이루어져야 한다.

2.2. 관련 모델 분석

현재 웹사이트 가시성에 영향을 주는 요인들에 대한 다양한 관련 연구가 진행 중이다. Binnedell 모델은 웹사이트 가시성에 관련된 초기의 분석 모델로서 어떠한 순위를 부여하지 않고 단순히 긍정적인 요인과 부정적인 요인으로 구분한다[5]. Chambers 모델은 Binnedell 모델의 한 단계 상승한 모델의 개념으로 동일한 속성들도 포함하나 실험을 통해 순위를 정한 모델이다[6]. 또한 Vissar 모델은 Chambers 모델을 기반으로 보다 세밀한 실험과 학문적 연구와 실무 전문가들의 면담을 통합하여 순위를 더욱 심층적으로 측정하였다[7].

<그림 2>에 나타난 Binnedell 모델에서 플러스 사인으로 나타난 요인들은 웹사이트에 있어서 긍정적 요인으로 나타나며 마이너스 사인은 부정적 영향을 주는 요인으로 구분한다. 실험적 경험이 이루어지지 않고 긍정적 요인이나 부정적 요인들을 동일하게 배열하여 순위를 결정하지 않은 것이 단점이나 이차적인 실험 결과에 따라 긍정적 요인과 부정적 요인으로 집단화 한 것으로도 웹 가시성 분석 모델로서 적합한 것으로 판정되었다[8]. 이 모델은 계층적 구분 없이 동일한 레벨로 6개의 항목이 부정적 요인으로 구분되었다[5].

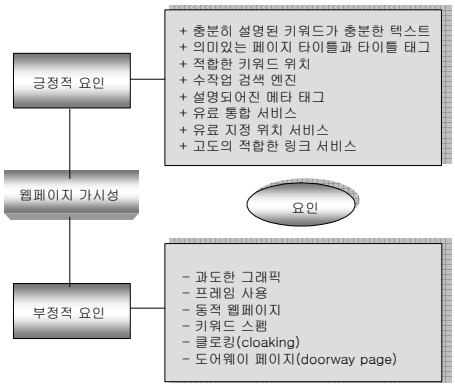

그림 2. Binnedell 모델
Fig. 2 Binnedell Model

Chambers 모델은 444개 웹사이트를 12개의 키워드로 6개의 검색엔진을 사용하여 실험을 하였으며 이를 통해 학문적 연구결과와 함께 동일한 분야의 전문가 심사를 통해 순위를 결정하여 모델을 완성하였다<그림 3>[6][8]. Binnedell 모델이 긍정적 요인과 부정적 요인으로 구분한 것과는 달리 점진적인 단계로 긍정적 및 부정적 요인들을 구분하여 완성된 것이 특징이다.

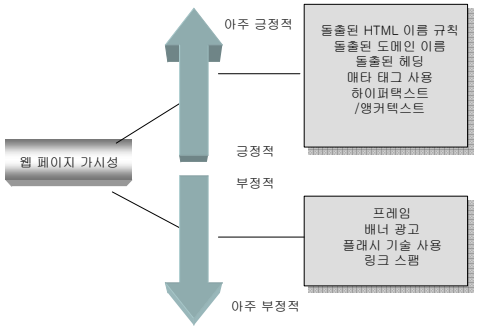

그림 3. Chambers 모델
Fig 3. Chamebers Model

Visser 모델은 Chambers모델과 동일하게 학문적 연구와 더불어 실험을 병행하였으며 실험결과를 긍정적 요인과 부정적 요인을 심층적으로 분석하여 필수, 부과, 경고, 위험 요인으로 부속 카테고리를 생성하여 모델을 완성하였다<그림 4>[7][8].

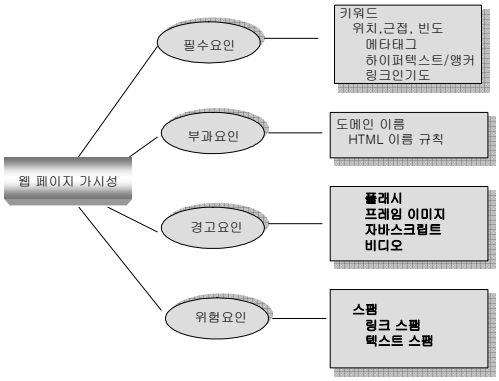

그림 4. Visser 모델
Fig. 4 Visser Model

2.3 분석 결과

<표 1>은 전 장에서 분석한 세 가지 모델들 중 부정적 요인들을 유사한 항목들 또는 동일한 항목들로 집합하여 분석한 결과이다.

표 1 적용 모델의 부정적 요인
Table 1. Negative Elements of Adaptive Models

Model	Binnedell	Chambers	Visser
그래픽	O		
도어웨이 페이지	O		
동적 웹페이지	O		
링크스팸		O	O
배너 광고		O	
자바 스크립트			O
클로킹	O		
키워드스팸	O		O
프레임	O	O	O
플래시		O	O

Binnedell, Chambers, Viseer 모델의 부정적 요인들을 유사한 항목이나 동일한 항목들을 통합하여 분석한 결과 프레임, 링크 스팸, 키워드(텍스트) 스팸, 및 프레임, 플레시 등 네 가지 부정적 요인들이 두 모델 이상에서 나타났다.

또한 <표 1>에 나타난 항목들의 연구 조사 결과를 종합적으로 분석하면 프레임은 일정한 항해 배치를 위해 웹사이트를 생성할 때 사용하나 일반적인 프레임으로 구성된 웹페이지는 검색 엔진 방문자들이 색인하거나 계속적인 링크로 연결되지

않는 것으로 분석되었다[9]. 프레임을 사용하지 않는 웹페이지들이 검색 색인의 방문율을 증가한다는 이유로 웹사이트를 설계할 때 프레임을 무시하거나 배제하는 경우가 많다[10]. 또한 일부 전문가들은 일반적인 웹페이지는 사용자들이 스크롤링을 회피하기 때문에 수직적인 스크린을 위한 프레임을 제한하여야 한다는 분석도 제기 되었다[11]. 키워드 스팸 색인의 경우에는 대부분 텍스트가 적합하게 읽혀질 수 있음에도 불구하고 중복적으로 키워드나 키워드 문장을 사용하는 경우에 나타난다. 해당 웹사이트를 소개하고 검색엔진 상위 순위를 유지하기 위해 방문자들에게 확신을 주기 위한 목적으로 키워드를 사용할 경우에는 웹사이트를 설계할 때 검색 엔진 색인에서 제거되지 않도록 하기 위해 텍스트는 시각적으로 중요하지 않는 부분에 위치하는 등의 기법을 사용하여 보이지 않도록 초점을 두는 것으로 해결할 수 있다[12]. 검색 엔진 알고리즘을 사용할 경우, 주어진 웹페이지의 방문되는 링크수를 증가시키기 위해 링크분석의 사용이 일반적이나 대부분이 다수의 의미가 없는 컨텐츠의 링크들로 이루어진다[13].

따라서 보다 효과적인 웹사이트 가시성을 고려한 웹 사이트를 설계하기 위해서는 다양한 방법론이 존재할 수 있으나 본 논문에서는 앞장에서 분석한 각 모델에서 결정된 요인들에 가중치를 부과하는 방식으로 순위를 결정하는 방법을 선택하였다.

III. 웹 가시성 순위 평가 실험

본 논문에서 순위를 결정하기 위한 실험은 앞 장에서 분석한 각 모델들에 점수를 배정하는 방식으로 결정하였다. 중복된 모델에서 동일한 부정적 요인들이 존재하나 각 모델들에서 부정적 요인들을 결정한 방법들이 다르기 때문에 이 모델들의 실험 결과를 바탕으로 서로 다른 가중치를 사용하는 방법을 선택하였다. SEO 분야의 기술들의 진화 과정에 따라 Chambers 모델에 임의의 점수를 부과하고 Binnedell 모델에 Chambers 모델 이전에 단순히 긍정적 요인과 부정적 요인으로 구분한 Binnedell 모델에 10점을 감점하고 Chambers 모델 이후에 보다 심층적으로 개발된 Visser 모델에 동일한 점수인 10점 가점을 배분한 Weideman 방식[8]을 채택하여 Chambers 모델에 임의의 30점을 배분하고 Binnedell 모델에 10점을 감점하며 Visser 모델에 10점을 가점하는 방식으로 배정하였다. 본 논문에서 제안된 모델들에 적용된 가중치들은 Weideman 방식을 기반으로 하였으며 가장 부정적인 항목의 가중치를 1로 채택하였다. 또한 부정적 요인들을 계량화하기 위하여 각 모델에 존재하는 레벨의 부정적 차이에 따라 가중치를 달리하였다[8].

3.1 Binnedell 모델

<그림 2>에서 Binnedell 모델의 경우, 여섯 개의 부정적 요인들을 동일하게 취급하였으므로 동일한 가중치를 부여하였다[8].

$$\text{Binnedell모델 평가점수} = \frac{\text{Binnedell모델 배정점수}}{\text{Binnedell모델 부정적요인수}}$$

표 2. Binnedell 모델 평가점수
Table 2. Binnedell Model Points

요인	레벨	가중치	평가점수
과도한 그래픽	부정적	1	3.3
프레임 사용	부정적	1	3.3
동적 웹페이지	부정적	1	3.3
키워드 스팸	부정적	1	3.3
클로킹	부정적	1	3.3
도어웨이 페이지	부정적	1	3.3
		가중치 총합: 6	평가점수총합:20

Chambers 모델의 경우, 부정적 요인들이 점진적인 네 가지 레벨로 분류되었으므로 각 레벨에 1씩 증가하여 가중치를 배정하였다<표3>[8].

$$\text{Chambers모델 평가점수} = \frac{\text{Chambers가중치} \times \text{Chambers배정점수}}{\text{Chambers모델가중치총합}}$$

표 3. Chambers모델 평가점수
Table 3. Chambers Model Points

요인	레벨	가중치	평가점수
가시적 링크스팸	아주 부정적	4	12
플래시 내용	부정적	3	9
프레임	비교적 부정적	2	6
배너광고	조금 부정적	1	3
		가중치 총합:10	평가점수총합:30

3.3 Visser Model

Visser 모델의 경우 일곱 개의 부정적 요인을 두 개의 부속카테고리로 분류한 것을 고려하여 가중치를 배정하였다<표4>[8].

$$\text{Visser모델 평가점수} = \frac{\text{Visser모델가중치 x Visser모델배정점수}}{\text{Visser모델가중치총합}}$$

표 4. Visser 모델 평가점수
Table 4. Visser Model Points

요인	레벨	가중치	평가점수
링크 스팸	경고	3	10.9
텍스트 스팸	경고	3	10.9
플래시	위험	1	3.6
프레임	위험	1	3.6
이미지	위험	1	3.6
자바스크립트	위험	1	3.6
비디오	위험	1	3.6
		가중치 총합:11	평가점수총합:40

3.4 평가 점수 적용

위의 분석 결과에서 부정적 요인들을 배제하기 위한 순위를 결정하기 위해 <표 3>의 각 모델에서 산출된 부정적 요인들의 평가점수들을 합산하여 점수가 높은 순위로 재배열하였다<표 5>. 즉 링크 스팸의 경우, Chambers 모델에 나타난 가시적 링크 스팸과 Visser 모델의 링크 스팸이 동일한 항목으로 분류되어 Chambers 모델의 가시적 링크 스팸의 평가점수 12점에 Visser 모델의 링크 스팸의 평가점수 10.9를 합산하여 <표 5>의 링크 스팸의 총평가점수는 21.9로 나타난다. <표 5>는 이러한 방식으로 <표 3>에 나타난 각 모델의 유사한 항목이나 동 일한 항목의 평가 점수를 합산하여 총평가점수가 높은 순위로 재배치한 분석 결과이다.

표 5. 웹사이트 가시성 부정적 요인 평가표
Table 5. Negative Element Point List for Website Visibility

Model	Binnedell	Chambers	Visser	총평가점수
링크 스팸		O	O	21.9
키워드 스팸	O		O	14.2
프레임	O	O	O	12.9
플래시	O		O	12.6
그래픽	O		O	6.9
자바스크립트			O	3.6
도어웨이 페이지	O			3.3
동적 웹페이지	O			3.3
클로킹	O			3.3
배너 광고		O		3.3

<표 5>에서 링크 스팸과 키워드(텍스트) 스팸이 상위 순위로 구분되어 웹사이트 가시성에서 가장 배제되어져야 할 부정적 요인으로 나타났다. 따라서 기업의 이익률을 극대화하기 위한 검색엔진 최적화를 위한 웹사이트를 설계할 때 우선적으로 배제되어져야 할 요인으로 고려될 수 있다.

IV. 결론

본 논문에서 제안한 검색엔진 최적화를 위한 웹사이트 가시성을 향상시키기 위한 실험 결과, 웹사이트에서 스팸 색인이 우선적으로 제거되어져야 하는 부정적 요인으로 나타났다. 본 논문에서 제안한 실험결과가 웹사이트 설계자들이 웹사이트를 설계할 때 검색엔진 최적화를 위해 웹 가시성을 향상시키기 위한 지침으로 제공될 수 있다. 본 연구에서는 학문적 연구에 중심을 둔 웹 가시성 분석모델을 기반으로 하였으나 향후에는 보다 실무적 관점의 요인들을 확대하여 적용할 뿐만 아니라 긍정적 요인들을 분석하여 웹사이트를 설계할 때 적극적으로 적용되어질 수 있는 지침을 마련할 수 있도록 연구가 확장되어져야 한다.

참고문헌

[1] M.Weideman, Website Visibility: The Theory and Practice of Improving Rankings, Chandos Publishing, Oxford, 2009

[2] R. Fishkin, Search Engine Ranking Factors V2, *http://www.seomoz.org/article/search-ranking-factors*, 2007

[3] 검색엔진최적화(SEO), 검색엔진마케팅, 유틸정보, *http://blog.naver.com/hunitra*

[4] 김성기, “멀티미디어 시스템에서 콘텐츠를 보호하기 위한 방법”, 한국컴퓨터정보학회 논문지, 제14권 제12호, 113-121쪽, 2009년, 7월

[5] M. Binnedell, The WBPN model: a proposed design approach to maximize whesite visibilityt osearch engines. Cape Technikon, 2003

[6] R. Chambers, Search engine strategies: a model to improve website visibility for SMME websites, CPUT, 2005

[7] E. Visser, Search engine optimization elements effect on website visibility, CPUT, 2007

[8] .Weideman, "Website Visibility- Quantifying Negative Search Engine Ranking Elements for Optimisation", Proceedings of Internet and Multimedia Applications, pp. 562-568, August, 2009

[9] J.Whalem, How to Optimize a Framed Site for High Ranking, *http://www.webproguide.com/article/*, 2000

[10] T. Dahm, Serach Engine Optimization Tips, *http://www.webdevelopersjournal.com/article/search_strategies_tips.html*, 2000

[11] E.Zeeger, Why Frames are bad for search engine optimization, *http://www.seomatrix.com/newsletter/archive/2006/SEO-Newsletter-July-2006/*, 2008

[12] M. Weideman, "Ethical issues on content distribution to digital consumers via paid placements as opposed to website visibility in search engine results", Proceedings of ETHICOMP pp. 169-186, 2004,

[13] A. Keyhanipour, "AGP-adaptive web ranking discovery framework based on combinative content and context feature, Journal of Informatics, No.3, Vol. 1, pp. 78-89, 2009

저 자 소 개

윤선희

1983 : 숭실대학교 공과대학
전자계산학과(학사)

1986 : (미)웨인주립대학교
전자계산학과(석사)

1999 : 성균관대학교 공과대학
전자계산전공(박사)

1986~1990 : (미),CSDC,PRC,
DUCOM, 시스템분석가

1991~1997 : KIST 시스템공학연구소
선임연구원

1997~2000 : ETRI 컴퓨터소프트웨어
기술연구소 선임연구원

2000~현재 : 숭의여자대학
인터넷정보과 교수

관심분야 : 유비쿼터스 컴퓨팅,
IT서비스, e-비즈니스