

웹로그 데이터 및 퍼지인식도를 이용한 웹쇼핑몰 사용자 구매패턴 분석

이건창

성균관대학교 경영학부 교수
(leekc@skku.ac.kr)

권순재

성균관대학교 경영학부 박사과정
(sjkwon@dragon.skku.ac.kr)

김재경

경희대학교 경영학부 교수
(jaek@khu.ac.kr)

.....

웹 쇼핑몰은 오늘날 디지털 경제를 상징하는 신경계의 중추적 역할을 하고 있다. 이러한 웹쇼핑몰이 가지고 있는 많은 장점 중의 하나는 사용자들의 구매행동이 웹로그 데이터로 자동으로 저장되어 이를 기초로 보다 구체적인 데이터 마이닝 작업을 할 수 있다는 점이다. 일반적으로 웹로그 자료는 자료의 방대함에 비하여 의미있는 정보를 추출하기가 상대적으로 난해하여, 데이터마이닝 작업시 많은 어려움을 겪는 분야이다. 본 논문에서는 이러한 웹로그 분석의 특징을 고려하여 웹로그 자료가 가지고 있는 인과관계를 효과적으로 추출하여 이로부터 사용자들이 가지고 있는 구매패턴을 추출할 수 있는 방법론을 제안한다. 본 논문에서 사용한 방법론은 퍼지인식도이며, 웹로그 자료는 실제로 운영 중인 B웹쇼핑몰에서 입수하여 분석하였다. 분석결과 본 논문에서 제안한 퍼지인식도 방법은 웹로그분석시 유용한 사용자 구매패턴을 추출할 수 있음이 확인되었다.

.....

1. 서 론

전세계 2억명 이상이 사용하는 인터넷은 이제 우리의 일상생활에 깊숙이 침투해 있다. 그동안 오프라인 상에서 구매에 익숙해져있던 소비자들도 이제 점차적으로 온라인 구매에 적응하고 있다. 이처럼 인터넷상에서 구매행위가 늘어가면서 학계에서도 인터넷상에서의 소비자 행위에 대한 연구가 많이 이루어지고 있다. 그 이유는 인터넷이 제공하는 환경 자체가 전통적인 환경과 매우 다르기 때문에 사용자의 행위를 효과적으로 파악할 필요가 점점 더 커지고 있기 때문이다. 특히, 인터넷 사이트들이 개인화되고 맞춤화된 정보를 제공하기 위해서는 웹

사이트 사용자의 행위를 보다 구체적이고 전략적으로 연구할 필요가 대두되게 되었다(Moe and Fader, 2000; Borges and Levene, 1999). 이러한 배경 하에 최근에 급부상하고 있는 웹 로그파일 분석은 웹사이트를 방문한 사람들의 방문 데이터를 서버가 인식하여 누가 언제 무엇을 요청했고 가져갔는지, 웹 서버에 얼마나 많은 사람이 접속하였는지, 어디에서 접속하였는지, 어떠한 사이트를 자주 방문하는지, 가장 오래 보는 또는 가장 많이 보는 페이지는 무엇인지 등등의 방문자 기록이 정리된 파일을 분석하는 기법으로 그 중요성이 높아져가고 있다(Spiliopoulou et al., 2000; Moe and Fader, 2000; Borges and Levene, 1999; Cooley, et al., 1997). 일반적으로 고객의 구매

패턴을 분석하고 이를 의사결정에 이용하려는 연구는 전통적으로 마케팅에서 로짓함수를 이용하여 오프라인 시장에서의 고객의 구매자료를 분석하는 연구에서 시작되었다고 볼 수 있다. 최근에는 통계학과 인공지능 기법에 기반한 다양한 데이터마이닝 기술이 개발되면서 이를 활용하여 고객의 특성 및 성향을 분석하여 구매 패턴을 분석하는 연구가 활발히 진행되고 있다(Chen, et al., 1996, 1998; Lawrence & Gile, 2000; Srivastava et al., 1999; Mobasher et al., 1999; Borges & Levene, 1999; Cooley, et al., 1999).

더욱이 인터넷 시대로 개편되면서 온라인 쇼핑몰을 통한 매출이 급증하게 되고 이들에 대한 관심이 높아지자 웹로그를 이용하여 인터넷 쇼핑몰을 이용하는 고객들의 구매 패턴을 분석, 활용하려는 연구가 시작되고 있다. 하지만 기존의 웹쇼핑몰에서 사용자의 구매패턴을 분석하고자 한 연구에서는 웹로그에 내재되어 있는 퍼지하고 불확실한 내용을 반영하지 못하였다는 한계점이 있다. 일반적으로 온라인 쇼핑몰에서의 사용자의 구매패턴은 모호하며 부정확한 특징을 가지고 있으므로 이를 적절히 표현하는 방법론이 필요하다. 퍼지인식도(Fuzzy Cognitive Map)는 이러한 부정확하고 모호한 개념을 표현하는데 적합한 특징을 가지고 있다.

이에 본 연구에서는 퍼지인식도를 활용하여 온라인 쇼핑몰의 사용자의 웹로그를 분석한 결과를 바탕으로 온라인 쇼핑몰 사용자의 구매패턴을 분석하고 이를 활용하여 적절한 온라인 쇼핑몰 구축전략을 제시하고자 한다. 이를 위하여 먼저 웹로그 자료를 분석하여 주요 요인을 도출하였으며, 데이터마이닝 기법 중에서 요인들간의 인과관계를 추출하는 기법을 이용하여 이들 요인들간의 규칙을 추출하여 퍼지인식도를 작성하였다. 이렇게 작성된 퍼

지인식도는 웹쇼핑몰에서 실제사용자들의 실제적인 구매패턴이 반영되어 있다는 장점이 있다. 이때의 퍼지인식도를 본 연구에서는 핵심 퍼지인식도라고 명하였다. 또한 웹쇼핑몰에서 사용자의 구매패턴의 긍정적인 요인을 반영하기 위하여 기존문헌을 분석하여 웹사이트를 구성하는 중요요인을 도출하고 이를 중심으로 전문가들에게 문의하여 퍼지인식도를 작성하였는데 이를 슈퍼 퍼지인식도라고 명하였다. 이상의 두 퍼지인식도를 통합하여 웹쇼핑몰에서 사용자의 구매 패턴을 표현하는 통합 퍼지인식도를 작성하였으며, 이렇게 작성된 통합 퍼지인식도는 웹쇼핑몰에서 사용자들의 다양한 구매패턴을 도식화하는데 사용된다. 또한 다양한 시뮬레이션을 통하여 실제 웹쇼핑몰에서의 주요 요인을 변화함으로써 사용자의 구매패턴에 미치는 영향 및 인과관계에 대한 분석을 실시하였다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 웹마이닝, 구매자 패턴, 웹쇼핑몰 사이트 디자인, 퍼지인식도에 대한 기존연구를 통하여 본 연구의 의의를 설명한다. 3장에서는 연구모형을 통하여 본 연구의 연구방법론을 소개하면서, 실제 쇼핑몰의 웹로그를 분석하고 퍼지인식도를 작성한다. 4장에서는 도출된 퍼지인식도를 이용하여 시뮬레이션을 실시하고 본 연구가 갖는 실무적인 시사점을 고찰한다. 마지막으로 5장에서는 향후 연구방향 및 결론을 제시한다.

II. 연구배경

본 절에서는 웹로그분석, 퍼지인식도 그리고 웹쇼핑몰 사이트를 구축하기 위한 기존연구를 고찰한다. 마지막으로 전통적인 상거래에서 사용자의 구매 패턴을 분석하는 마케팅의 연구를 고찰함으로써 본 연구의 이론적인 바탕을 설명하고자 한다.

2.1 웹 로그분석에 관한 기존연구

사용자가 웹사이트를 이용하면 사용자가 해당 웹사이트에서 움직인 결과가 웹로그(web log)라는 형태로 기록된다. 따라서 웹로그분석이란 로그데이터를 기반으로 다양한 정보를 추출해 내는 것으로 정의할 수 있다(Cooley, et al., 1997). 그런데 웹로그분석은 분석 목적에 따라 단지 로그 정보를 분석하는 것에 한정시키기도 하고 로그 정보를 기반으로 한 보다 다양한 정보를 분석하는 확장된 개념으로 확대시키기도 한다. 일반적 의미의 웹로그 분석은 로그 데이터를 이용하여 트래픽을 파악하고, 이 트래픽이 지닌 의미를 분석해 나가는 것이라고 할 수 있다. 로그 데이터를 이용하여, 웹사이트의 페이지뷰, 사용자별 페이지뷰, 접속장소 및 방식, 시간별 페이지뷰, 방문자수 등에 대한 현황 및 추세를 분석하는 것이다. 웹사이트의 클릭흐름을 분석하는 것 역시 이 범주에 들어간다. 사용자가 웹사이트를 방문하는 경로와 서핑하는 경로에 대한 분석을 통하여 웹사이트가 지닌 문제점을 찾고, 사용자가 웹사이트에서 무엇을 원하는 지를 보다 구체적으로 파악하는 것이다. 그러나 이러한 분석은 웹 로그에 대한 1차원적인 정보외에는 추출이 불가능하기 때문에 웹 로그가 가지고 있는 최소

한의 정보밖에는 추출할 수가 없다. 이에 반해, 확장된 의미의 웹로그 분석은 단지 로그 데이터뿐 아니라, 웹사이트에서 보유하고 있는 고객등록정보, 구매정보, 외부환경정보 등을 복합적으로 사용하는 분석을 말한다. 이러한 분석을 통하여 사용자 특성별로 웹사이트의 이용, 구매에 대한 보다 폭넓은 분석이 가능해진다. 본 연구에서 언급하는 웹로그 분석의 개념은 후자에 속한다고 볼 수 있다.

특히 웹 로그 분석은 최근에 새로운 화두로 떠오른 데이터마이닝 개념과 결합되어 웹 마이닝으로 불리기도 하며, 컴퓨터 사이언스 분야와 마케팅 분야, 통계분야 등에서 많은 연구가 진행되어 왔다. 그러나, 웹 마이닝에 사용되는 로그파일의 경우 그 크기가 매우 방대하고 전처리(Pre-Processing) 과정을 거치기 전에는 원시자료(Raw Data) 만으로는 분석이 불가능한 관계로 일반 사용자들이 다루기는 많은 한계점을 가지고 있다. 특히 실제 웹 로그 파일 중 전처리 과정을 거치면 10%이내의 자료만이 분석에 사용되기 때문에 일반 사용자들이 웹 마이닝을 시도하기란 매우 어려운 실정이다.

웹로그 분석에 대한 기존연구는 아직 시작단계에 있으므로 많은 연구가 이루어진 것은 아니지만, 그동안 이루어진 연구를 고찰해보면 웹데이터에서 지식을 추출하기 위하여 규칙이나 패턴을 발견하는 알고리즘을 개발하는데 중점을 둔 컴퓨터 과학분야의 연구와 이를 이용하여 전략적인 의미를 도출하려는 응용분야의 연구로 나뉘어질 수 있다.

웹데이터에서 지식을 추출하기 위하여 규칙이나 패턴을 찾아내기 위한 연구의 경우 알고리즘을 중심으로 연관규칙 분석, 군집 및 분류규칙, 순차적인 패턴 세 부분으로 이루어져 왔다. 이중 이번 연구에서 사용되며, 현재 웹마이닝에서 가장 많이 쓰이는 연관규칙 분석을 중심으로 기존문헌을 통하여

살펴보기로 한다. 연관규칙 분석은 기존의 데이터 마이닝 연구분야 중에서 가장 활발한 연구가 진행된 분야로써 항목간에 연관관계를 중심으로 규칙을 탐색하는 분야이다(Agrawal and Srikant, 1994 (a,b); Park et al., 1995; Savasere et al., 1995). 이 분야는 대규모의 슈퍼마켓 등에서 판매한 물품들의 상호연관성을 찾아내어 사용자가 미리 정한 지지도와 신뢰도를 바탕으로 연관 규칙을 찾아내는 것이다. 예를 들어 기저귀를 사는 사람이 맥주를 동시에 구매하는지를 알고자한다고 가정하자. 이를 연관규칙을 이용하여 일정한 규칙을 추출해보니 “기저귀→맥주”(지지도=15%, 신뢰도=75%)로 분석되었다. 여기서 지지도 15%의 의미는 전체 고객 중에서 15%가 기저귀와 맥주를 동시에 구매한다는 것이고, 신뢰도 75%의 의미는 기저귀를 구매한 고객 중에서 75%만이 맥주를 구매한다는 것이다. 지지도는 전체 구매 고객 중에서 연관 규칙에 관련된 항목들을 구매하는 정도를 나타내고, 신뢰도는 특정 물품과 물품 사이의 연관성을 나타내므로, 일반적으로 수많은 추출된 규칙 중에서 먼저 최소한 주어진 지지도 이상의 물품들을 찾아내고 그 다음으로 물품들 사이의 신뢰도를 측정하게 된다. 슈퍼마켓 등에서 판매된 물품들의 판매트랜잭션으로 구성된 데이터베이스를 “바스켓 데이터베이스”(Basket-Database)라 부르고, 초기의 연관 규칙 탐색 알고리즘은 이 분야에서 연구가 시작되었다(Srikant and Agrawal, 1995).

연관 규칙을 탐색하는 방법은 크게 두 가지 부분으로 나누어진다. 첫째, 최소 지지도 이상으로 발생하는 항목을 찾아내는 방법과 둘째, 이들 사이의 신뢰도를 측정하여 한계치(threshold) 이상의 값

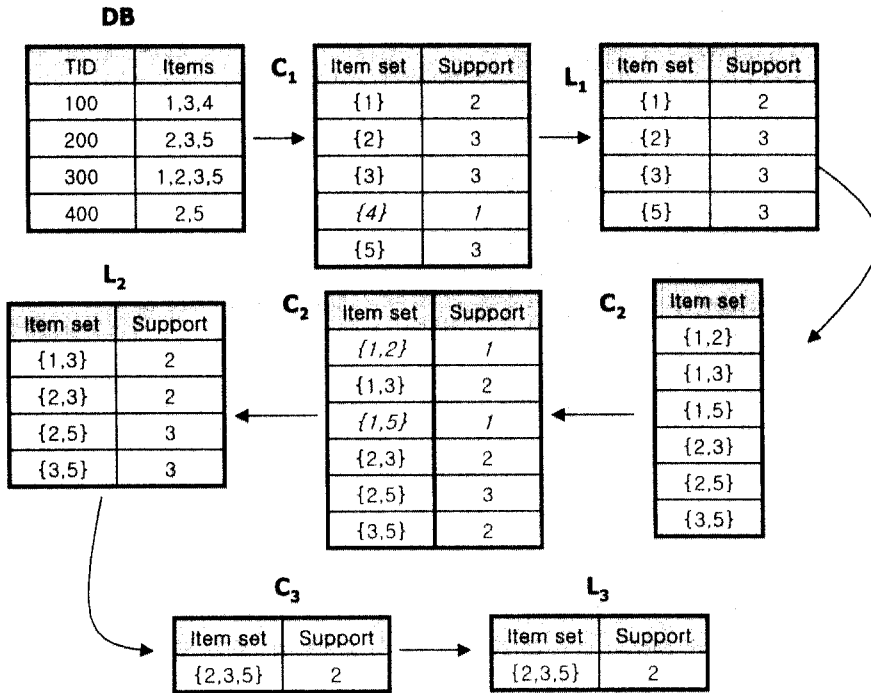
을 보일 때만 연관 규칙을 만들어내는 것이다. 이를 효율적으로 해결하고자하는 연구가 Agrawal이 제안한 Apriori 알고리즘(Agrawal and Srikant, 1994ab; Agrawal et al., 1993ab)이며, 이를 개선한 것이 DHP이다(Park et al., 1995; 박종수, 1996). 연관분석의 대표적인 Apriori 알고리즘을 <그림 1>을 통하여 구체적으로 살펴보자.

거래가 위의 <그림 1>의 데이터베이스 처럼 발생했다고 가정하자. 이때 첫 번째 거래(TID=100)에서는 item 1, 3, 4가 팔렸다는 의미이다. 그러면 먼저 항목(Item)이 하나인 후보집합(Candidate Set)의 경우를 계산하면 C1과 같다. 이를 지지도(Support)가 2이상인 경우만을 추출하면 항목이 “4”인 경우에는 지지도가 1이므로 제외되고 지지도가 1이상인 거대항목집합(Large Item-set)인 L1이 생성된다. 이를 기준으로 이번에는 항목(Item)이 둘인 경우를 계산하여 C2와 같은 후보집합을 생성하고 여기서도 지지도가 2이상인 항목({1,2}, {1,5})을 제외하고 L2를 생성하게 된다. 이를 기준으로 다시 항목(Item)이 셋인 경우를 계산하여 C3와 같은 후보집합을 생성한다. 이때 L2를 이용하여 항목수가 세 개인 후보집합은 {1,2,3}, {2,3,5}, {1,3,5} 세 개이지만 이중에서 {1,2,3}과 {1,3,5}는 앞선 C2 계산과정에서 {1,2}, {1,5}가 포함되지 않았으므로 후보집합에서 제외된다. 따라서 최종적으로 남는 거대항목집합은 {2,3,5}가 되고 이를 바탕으로 도출된 연관규칙을 생성하게 된다. 예를 들어 거대항목집합인 {2,3,5}에서 도출한 2→3인 연관규칙의 경우 지지도는 50%이며, 신뢰도는 100%로 설명된다(Charu, et al., 1998).

이밖에도 군집 및 분류 규칙을 이용한 연구와 순

1) 이를 장바구니 분석이라고 명하기도 한다.

〈그림 1〉 Apriori 알고리즘 예제



차적인 패턴을 이용한 연구 등도 웹마이닝에서 이루어지고 있다. 군집 및 분류규칙을 이용한 대표적인 연구로써 Igor(2000)등은 웹에서 사용자들의 거래자료의 형태가 관계형 데이터 베이스 형태로 일정하게 고정되어 있지 않은 경우 이를 처리할 수 있는 확률기반의 군집 프레임워크를 제안하였다. 또한 Fu(1999)등은 웹 사용자의 접속 패턴을 클러스터링하고 이를 통하여 일정한 규칙을 제공하는 방법론을 개발하였으며, McCallum(2000) 등은 사용자들의 웹데이터를 다차원의 데이터집합으로 관리하고 있는 경우 이를 효율적으로 군집화할 수 있는 방법론을 제시하였다. 순차적인 패턴이란 일정한 시간동안 시간에 따라 순차적으로 발생하는 거래로 정의할 수 있는데(Srivastava et al.,

1999), 이는 비교적 다양한 방법으로 이루어져 왔다. 대표적인 연구자로서 Agrawal와 Srikant (1994 a,b)는 순차적인 고객의 거래자료에서 일정한 패턴을 인식하는 알고리즘을 제시하였으며, 이후 Mannila 와 Meek(2000)는 웹에서의 사용자들의 거래 자료를 순차적인 패턴으로 인식하는 방법론을 개발하였다. 또한 Mannila(1997) 등은 순차적으로 발생하는 일련의 사건을 부분적인 집합으로 함께 묶어 에피소드(episodes)라 새롭게 정의하고 이를 기본으로 연관규칙을 도출하는 것이 기존의 다른 방법보다 그 효율성이 우수함을 증명하였다.

반면 이를 이용하여 전략적인 의미를 도출하려는 응용분야의 연구는 크게 웹사이트의 콘텐츠의 특성

을 분류하고 이를 전략적으로 활용하려는 연구방향과 웹에서 사용자의 패턴을 분석하여 이들의 연관관계를 분석하고 이를 전략적으로 활용하려는 연구방향으로 나눌 수 있다.

첫째, 웹사이트의 콘텐츠의 특성을 분류하고 이를 중심으로 전략을 제시하기 위한 연구를 살펴보면, Lawrence 와 Gile(2000)는 웹에서 존재하는 커뮤니티(Community)의 효율성을 판별하기 위하여 몇몇 커뮤니티의 사이트에 가중치를 부여하고 이를 검증함으로써 효율적으로 커뮤니티를 이용하는 전략을 제시하였다. Cooley(1999)등은 웹사용자들의 데이터를 사용자(Usage), 콘텐츠(Contents), 구조(Structure)의 세 가지로 분류하고, 사용자의 패턴을 분석하여 가중치가 높은 웹페이지를 전략적으로 활용하는 방안을 제시하였다. 이밖에도 Spiliopoulou(2000)등은 웹사이트의 효과성(Effectiveness) 향상시키기 위해서는 웹데이터마이닝을 통하여 사용자의 특성을 분석하는 과정이 필요함을 주장하였다.

둘째, 웹에서 사용자의 패턴을 분석하여 이들의 연관관계를 분석하고 이를 전략적으로 활용하려는 연구도 이루어지고 있다. 대표적인 연구로써 Srivastava(1999)등은 웹데이터에서 사용자의 패턴을 분석하고 이를 통하여 일정한 연관규칙과 지식을 획득하는 분석과정을 제시하였다. 특히 이들은 웹 사용자들의 패턴을 분석하는 과정에서 개인화와 사용자의 특성에 중점을 두고 차별화된 웹페이지 전략을 제안하였다. Mobasher(1999)등은 사용자가 자주 방문하는 사이트를 중심으로 이들의 연관관계를 분석하는 새로운 메커니즘을 제시하였다. 그 결과 자주 방문하는 사이트에 점수를 부여하여 많이 방문하지 않는 사이트를 적극적으로 추천하는 전략을 제시하였다. Borges와 Levene(1999)

는 하이퍼텍스트로 연결되어 있는 사이트의 웹페이지를 중심으로 사용자의 항해 패턴을 분석함으로써 웹페이지 디자인에 대한 몇가지 전략을 제시하고 있다. 또한 앞서 설명한 Cooley(1999) 등 역시 로짓함수를 이용하여 사용자의 패턴을 분석하였다. 이상의 기존 연구를 표로 요약하면 다음 <표 1>과 같다.

2.2 퍼지인식도에 관한 연구

Kosko(1986)에 의해 처음 소개된 퍼지인식도(Fuzzy Cognitive Map)는 인과관계값을 방향뿐만 아니라 변화의 크기까지 정하여 표현한 것을 의미한다. 결국 퍼지인식도는 기존의 인식도(Cognitive Map: CM)를 더욱 정교하고 더욱 풍부한 정보로 표현한 것이라고 할 수 있으며, 영향관계가 파악이 불가능할 때의 영향관계의 불확실성을 제거할 수 있는 기법으로 알려지고 있다(Kardaras & Karakostas, 1999). 이러한 퍼지인식도는 이와 관련된 다양한 연구를 나았는데 Montazemi & Conrath(1986)은 정보시스템 요구사항을 분석하기 위하여 퍼지인식도를 사용하였다. 이들은 정보시스템 전문가들을 대상으로 정보시스템 요구사항에 필요한 핵심 속성을 도출하고 이를 설문지를 통하여 인과관계를 추출하였다. 이 연구에서는 정보시스템 요구사항 분석을 위한 퍼지인식도의 구체적인 추론과정이 자세히 소개되고 있지는 않지만 설문지를 이용한 퍼지인식도 작성이라는 새로운 방법론을 제시하고 있다. Kim & Lee(1998)은 퍼지인식도가 인과관계 지식을 영역으로 하고 분야의 지식을 효과적으로 획득하여 전문가시스템을 구축할 수 있게 하는 유용한 도구이나 퍼지인식도의 퍼지관계(Fuzzy Implication)을 잘못 이해할 경우

〈표 1〉 웹로그 분석 기존연구

분야	분류	연구내용	연구자
컴퓨터 과학분야	연관규칙	항목(Item)간에 존재하는 연관관계 규칙을 탐색	Agrawal and Srikant(1994a); Parket al.(1995); Savasere et al.(1995);
	군집 및 분류	화물기반의 군집 프레임워크를 이용하여 고객들의 웹사이트 방문 형태를 다양하게 분류	Igor et al.(2000)
		웹 사용자의 접속 패턴을 클러스터링하고 이를 통하여 일정한 규칙을 제공	Fu et al.(1999)
		다차원 데이터베이스를 이용하는 사용자들의 웹로그 자료를 효율적으로 군집화할 수 있는 방법론 제시	McCallum et al.(2000)
	순차적 패턴	순차적인 거래자료에서 일정한 패턴을 인식하는 알고리즘을 제시	Agrawal & Srikant(1994b)
		웹 자료를 순차적인 패턴으로 인식하는 방법론을 개발	Mannila & Meek(2000) Srikant & Srikant(1996)
		순차적으로 발생하는 일련의 사건을 에피소드(episodes)라 새롭게 정의 및 분석	Mannila et al.(1997)
의사결정 응용분야	웹사이트의 콘텐츠 특성을 연구	커뮤니티의 효율성을 판별하기 위하여 몇몇 커뮤니티 사이트에 가중치를 부여하고 이를 검증.	Lawrence & Gile(2000)
		웹 사용자에 관한 자료를 사용(Usage), 콘텐츠(Contents), 구조(Structure)의 세 가지로 분류	Cooley et al.(1999)
		웹사이트의 효과성을 향상시키기 위하여 사용자의 특성을 분석	Spiliopoulou et al.(2000)
	웹 사용자의 패턴을 분석	특정 웹사이트에서의 사용자의 이용패턴을 분석하고 이를 통하여 개인화 작업을 수행	Srivastava et al.(1999)
		자주 방문하는 사이트를 중심으로 이들의 연관관계를 분석하고 자주 방문하는 사이트에 점수를 부여하여 많이 방문하지 않는 사이트를 추천	Mobasher et al.(1999) Borges & Levene(1999)
		로짓함수를 이용하여 웹 사용자의 이용패턴을 분석	Cooley et al.(1999)

잘못된 퍼지인식도를 얻을 수 있다고 주장하였다. 이에 이들은 퍼지인식도의 인과관계를 표시하는 퍼지인과관계와 퍼지부분인과관계를 포함하는 새로운 퍼지관계를 정립하고 이를 논리적으로 증명하였다. Banini & Bearman(1998)은 미네랄 산업에서

슬러리 유동학에 영향을 미치는 요인을 세명의 전문가들로부터 추출하고 이를 결합하여 사용하는 활용에 대하여 소개하였다. 이들은 전문가의 지식을 결합할 때 Max 연산을 취하는 방법을 사용하였는데 분석결과 미네랄 산업에 유용한 정보를 얻을 수

있었다. Schneider et al.(1998)은 사용자가 제공한 자료를 바탕으로 자동적으로 퍼지인식도를 구축하는 방법론을 제안하였다. 이 방법론은 벡터로 표현되는 두 변수간의 유사 정도(Similarity Degree)를 파악할 수 있도록 되어 있는데 영향관계의 방향성도 같이 분석이 가능하기 때문에 퍼지전문가시스템 셀에 사용될 수 있다고 강조하였다. 이들은 인구통계-경제 자료를 바탕으로 이 방법론의 유용성을 검증하였다.

Kardaras & Karakostas(1999)는 정보시스템 전략계획분야에 퍼지인식도를 적용하였다. 이들은 새로운 IT 프로젝트를 평가하기 위한 수많은 새로운 모델들이 출현하고 있지만 연구자들은 주로 경영분야 혹은 IT분야 한쪽에 국한하여 연구를 수행하고 있다고 주장하고 퍼지인식도를 이용하여 SISP 과정을 시뮬레이션 함으로써 대안적인 모델링 접근 방법을 소개하고자 하였다. 이들은 연구에서 경영과 IT영역을 포괄하는 165개의 변수와 210개의 인과관계를 갖는 대형 퍼지인식도를 사용하여 연구를 수행하였다. Satur & Liu(1999)는 정보시스템의 구축에서 정량적 또는 정성적인 자료의 확보가 용이하지 않음을 지적하고 퍼지인식도를 이용한 정보시스템 구현의 가능성에 대하여 소개하였다. 이들은 특히 지리정보시스템을 예로 들어서 퍼지인식도를 사용할 경우 지리적 특성간의 관계를 바탕으로 보다 일반화된 정보를 제공해 줄 수 있다고 주장하였다. Lee & Han(2000)은 EDI 통제에 퍼지인식도를 이용하여 EDI 통제에 영향을 미치는 내부요인, 외부요인, 그리고 자동화 요인간의 영향관계를 분석할 수 있음을 보여주었다. 이들은 EDI 통제 요인간의 영향관계를 파악하기 위해서 구조동식모델을 사용하였는데 이때 요인간의 인과관계가 중치를 퍼지인식도의 인과관계 값으로 사용하는

EDI 퍼지인식도를 제안하였다. 이들은 이 모형을 통해서 EDI 감사자가 EDI 통제에 대한 인과관계에 대해 보다 폭 넓은 지식을 얻을 수 있을 것이라고 하였다. Kwahk & Kim(1999)은 기업의 BPR에 퍼지인식도를 적용하였는데, 기업의 BPR 구축시 두 단계의 퍼지인식도를 이용하여 BPR시 기업에서 발생할 수 있는 갈등 및 다양한 비즈니스 문제에 대한 영향 요인을 분석하였다.

이외에도 퍼지인식도의 속성을 이용하여 퍼지인식도를 시뮬레이션(Fu, 1991), 조직적 전략모델링(Paradice, 1992), 문제정형화 및 의사결정분석지원(Warren, 1995; Heintz & Acar, 1992; Diffenbach, 1982; Fiol, 1992), 지식베이스 구축(Taber, 1991; Nakamura et al., 1982), 가상공간에서의 행동 모델링(Dickerson & Kosko, 1994), 분산협동 에이전트의 조정(Zhang et al., 1992), 주식시장 분석(Lee & Kim, 1997)과 같은 다양한 의사결정 문제에서 활용할 수 있도록 한 연구들이 있다. 이와같이 퍼지인식도는 문제를 구성하는 개념들간에 복잡한 인과관계가 존재하는 비구조적이고 본질적으로 퍼지한 문제영역에서의 지식습득 및 추론에 매우 유용한 것으로 알려져 있다. 아울러 퍼지인식도는 그 성질상 인과관계 지식을 행렬의 형태로 표현하기 때문에 퍼지인식도에 기초한 추론과정이 일정한 횟수에서 수렴이 되는지의 여부를 수리적으로 계산할 수 있는 장점이 있다(Taber, 1991).

본 연구의 경우 웹로그 분석을 통하여 사용자의 패턴을 반영하고자 한다. 일반적으로 웹로그에서 제공하는 정보는 레코드의 수는 매우 많으나 상대적으로 필드의 수가 적으므로 실질적으로 의사결정에 유용한 정보를 가지고 있지 못한 경우가 많다. 따라서 웹로그가 내포하고 있는 사용자 패턴에 대

〈표 2〉 최근 퍼지인식도에 관한 연구

연구분야	연구자	연구내용 및 공헌도	특징
퍼지인식도 활용	Montazemi & Conrath(1986)	정보시스템 전문가들을 대상으로 정보시스템 요구사항에 필요한 핵심 속성을 도출하고 이를 설문지를 통하여 인과관계를 추출	설문지를 사용하여 인과관계 파악
	Banini & Bearman(1998)	미네랄 산업에서 슬러리 유동학에 영향을 미치는 요인을 세명의 전문가들로부터 추출하고 이를 결합 및 활용 방안 소개	미네랄 산업에 적용 (Max 연산을 통한 지식 결합)
	Kardaras & Karakostas(1999)	퍼지인식도를 이용하여 정보시스템 전략계획 과정을 시뮬레이션 함으로써 대안적인 모델링 접근방법을 소개	SISP 시뮬레이션에 응용
	Satur & Kiu(1999)	정보시스템의 구축에서 정량적 또는 정성적인 자료의 확보가 용이하지 않음을 지적하고 퍼지인식도를 이용한 정보시스템 구현의 가능성에 대하여 소개	GIS에 응용
	Lee & Han(2000)	EDI 통제에 퍼지인식도를 이용하여 EDI 통제에 영향을 미치는 내부요인, 외부요인, 그리고 자동화 요인간의 영향관계를 분석	EDI 통제에 적용
	Kwahk & Kim(1999)	기업의 BPR 구축에 두 단계의 퍼지인식도를 이용하여 BPR시 기업에서 발생할 수 있는 갈등 및 다양한 비즈니스 문제에 대한 영향 요인을 분석함.	BPR 구축에 적용
인과관계지식 획득	Kim & Lee(1998)	퍼지인식도의 인과관계를 표시하는 퍼지인과관계와 퍼지부분인과관계를 포함하는 새로운 퍼지관계를 정립	논리적 검증
퍼지인식도 구축	Schneider et al.(1998)	사용자가 제공한 자료를 바탕으로 자동적으로 퍼지인식도를 구축하는 방법론을 제안	인구통계-경제자료 사용

한 정보는 매우 비구조적이며 불확실한 경우가 많다. 퍼지인식도의 경우 이 같은 비구조적이며 불확실한 정보를 반영할 수 있으므로 보다 나은 의사결정에 도움을 줄 수 있다.

2.3 웹 쇼핑몰 디자인에 대한 연구

웹 쇼핑몰 사이트 디자인의 우수성 여부가 매출액에 직접적으로 영향을 미치는 인터넷 쇼핑몰의 경우 성공적인 인터넷 사업을 위해 웹 쇼핑몰 사이트 디자인의 중요성이 강조되면서 이에 대한 연구

는 매우 활발이 이루어지고 있다. 특히, 웹 쇼핑몰 사이트 디자인에 대한 국내외의 연구동향을 살펴보면 웹 쇼핑몰 사이트 디자인 분야와, 웹 쇼핑몰을 구축시 핵심성공요인을 연구하는 분야로 나눌 수 있다.

웹 쇼핑몰의 디자인 분야의 경우 인간의 감성 분야를 연구하는 감성공학을 중심으로 발전하였다. 대표적인 연구자로서 김진우 & 문재운(1997)이 고객의 특정 감성을 유발할 수 있는 전자상거래 시스템의 사용자 인터페이스의 설계방안을 소개하였다. 연구결과 사용자 인터페이스 내의 특정 시각

디자인 요소의 조작을 통해서 신뢰감과 같은 고객의 감성 유발에 영향을 미칠 수 있음을 지적하였다. Lightfoot(1997)은 웹 사이트의 구조적인 측면에서 기존의 웹 사이트 설계 방법이 갖는 문제점을 웹 마스터와 최종사용자의 관점에서 지적하고 하이퍼텍스트와 시각화(Visualization) 소프트웨어를 이용하여 전통적인 웹 사이트 디자인 방법론이 갖는 문제점을 극복할 수 있는 방안을 제시하였다. Larson & Czerwinski(1998)은 웹 사이트의 메뉴를 디자인하는데 있어서 깊이(Depth)와 넓이(Breadth) 간의 균형의 중요성을 언급하고 깊이와 넓이가 서로 다른 웹 사이트를 만들어 웹 사이트 구조의 탐색 성과(Search Performance)를 파악하고자 하였다. De Troyer & Leune(1998)는 기존의 웹 사이트가 사용자 중심이 아닌 웹을 만드는 사람의 입장에서 구축되었기 때문에 많은 문제점을 내포하고 있다고 지적하고 웹이 가지고 있는 '자료 중심적'인 특성이 아닌 '사용자 중심적'인 방법으로 웹 사이트를 디자인하는 방법을 제안하였다.

이상의 웹사이트 디자인에 관한 연구와 더불어 웹쇼핑몰 사이트를 구축할 때 필요한 핵심성공요인에 대한 연구도 함께 이루어졌다. 대표적인 연구자로서 김광용 & 김기수(1999)는 인터넷 설문조사를 활용한 사이버 쇼핑물 디자인에 관한 연구에서 237명을 E-mail을 통해 설문조사를 실시하였다. 안준모 & 한상록(2000)은 성공적인 인터넷 쇼핑물을 구축하는데 있어서 해당 웹 사이트의 구성디자인의 특성과 마케팅 기능이 중요함을 강조하였다. 이들은 인터넷 상에서 제품구입 경험자 또는 인터넷 쇼핑물 전반적인 이용만족에 영향을 주는 인터넷 쇼핑물 구성기능상의 특징과 마케팅 기능요소를 제시하였다. Lohse & Spiller(1998)은

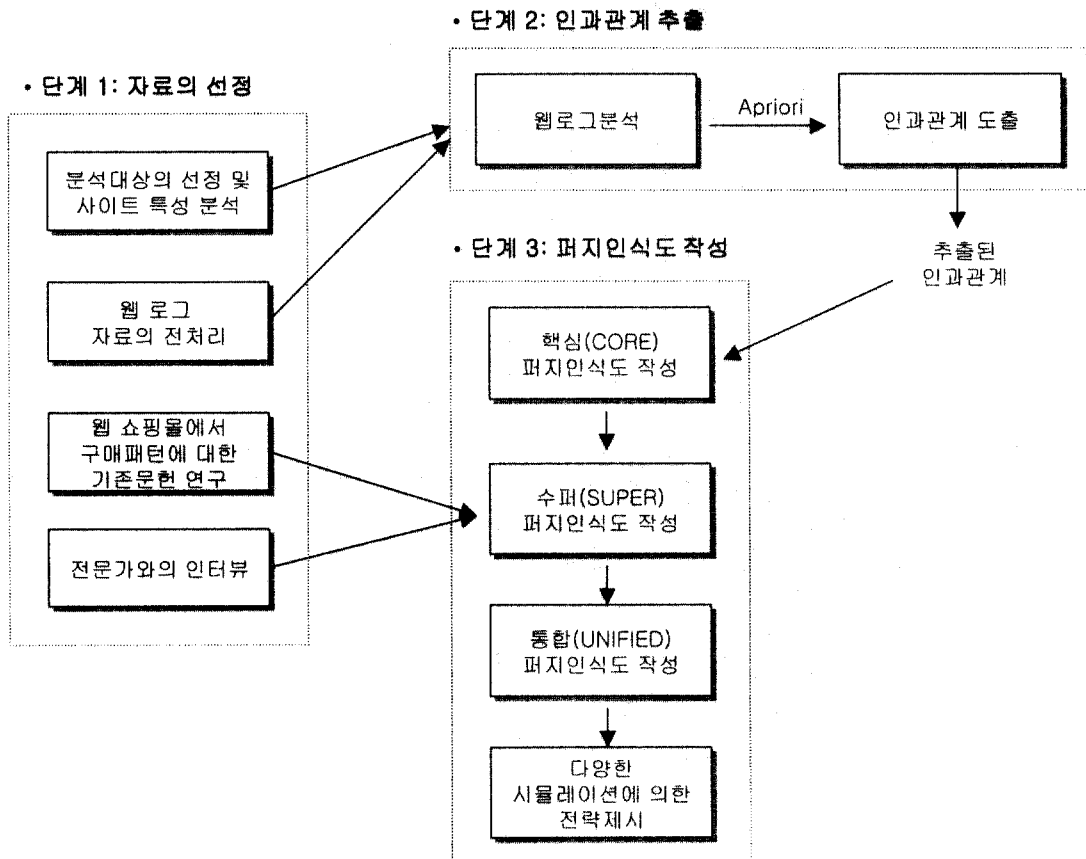
인터넷 쇼핑물의 링크의 수, 이미지의 크기, 제품의 개수, 네비게이션 형태 등과 같은 사용자 인터페이스 디자인의 특성과 소비자의 실제 이동량(traffic)과 판매량과의 관계를 회귀분석을 통해서 실증적으로 검증하였다. 이 연구는 웹 사이트 디자인을 소비자의 행동과 연관시켜 분석한 초기의 연구로 그 연구 결과가 매우 흥미롭다. Huizingh(2000)은 웹 사이트를 구성하는 요인을 크게 내용과 디자인으로 구분하고 이를 구별하는 프레임워크를 제시하여 기업들이 얼마나 다른 목적으로 웹을 사용하는지 실증적으로 분석하였다. 분석결과 그는 웹 사이트의 내용과 디자인을 구분하는 것은 의미 있으며 모든 웹 사이트의 내용과 디자인의 형태는 해당 사이트의 크기와 매우 밀접함을 증명하였다. Zhang et al.(2000)은 허츠버그(Herzberg)의 2요인이론(Two Factor Theory) 혹은 위생-동기이론(Hygiene-Motivational Theory)을 이용하여 웹 사이트 디자인이 사용자의 만족도에 미치는 영향을 분석하고자 하였다. Liang and Lai(2000)은 인터넷 쇼핑물을 설계하는 구체적인 방법론으로 동기요인, 위생요인, 매체풍부성요인의 세가지 카테고리를 바탕으로 이들이 소비자의 구매의사결정에 어떠한 영향을 미치는지 실증분석하였다. Wells & Fuerst(2000)는 웹 인터페이스에서 인터페이스 은유 개념을 이용하는 것이 중요하다고 지적하고 인터페이스 모드와 정신적 모형 형태(Mental Model Type)로 구성된 연구를 실시하였다. 연구결과 컴퓨터-인간 상호작용에 관한 연구나 인지심리학의 분야에 매우 유용하게 사용될 수 있음을 증명하였다. Forbes & Rothschild(2000)은 한상린(1998)과 Novak & Hoffman(1997) 등이 이미 언급한 flow개념을 이용하여 인터넷 상에서의 소비자의 경험을 중심으로 웹 사이트를

디자인할 수 있는 방법론을 소개하였다. 여기서 flow란 최적의 경험을 유도하는 과정을 의미한다 (Csikszentmihalyi, 1990). 이 연구가 기존의 연구와 차별화되는 점은 소비자의 행동을 웹 사이트 디자인과 연결시켰다는 점인데 사용자의 컴퓨터에 대한 숙련도와 정보/서비스의 난이도에 따른 웹 사이트 디자인 방법을 제시하였다는 특징을 가지고 있다.

III. 연구모형 및 분석

본 연구에서는 웹로그 분석을 통한 인과관계를 바탕으로 퍼지인식도를 구축하고 이를 추론 시뮬레이션을 통하여 영향력의 변화를 통하여 실제 구매한 사용자들의 "구매행동"이라는 성과를 중심으로 분석하고자 다음 <그림 2>와 같은 3단계 연구모형을 사용하였다.

<그림 2> 연구모형



3.1 자료의 선정

웹 쇼핑몰에서의 사용자의 구매패턴을 분석하기 위한 지식을 추출하기 위하여 첫째, 분석 대상이 되는 쇼핑몰을 선정하고 이 사이트의 특성을 분석하는 것이다. 둘째, 선정된 분석대상의 웹로그를 전처리하여 데이터마이닝이 가능한 상태로 만드는 것이다. 셋째, 웹쇼핑몰에 대한 기존연구를 통하여 사용자의 구매패턴에 영향을 주는 주요한 요인을 도출하였으며, 이를 중심으로 전문가들의 의견을 수렴하여 이들의 인과관계를 확정하였다.

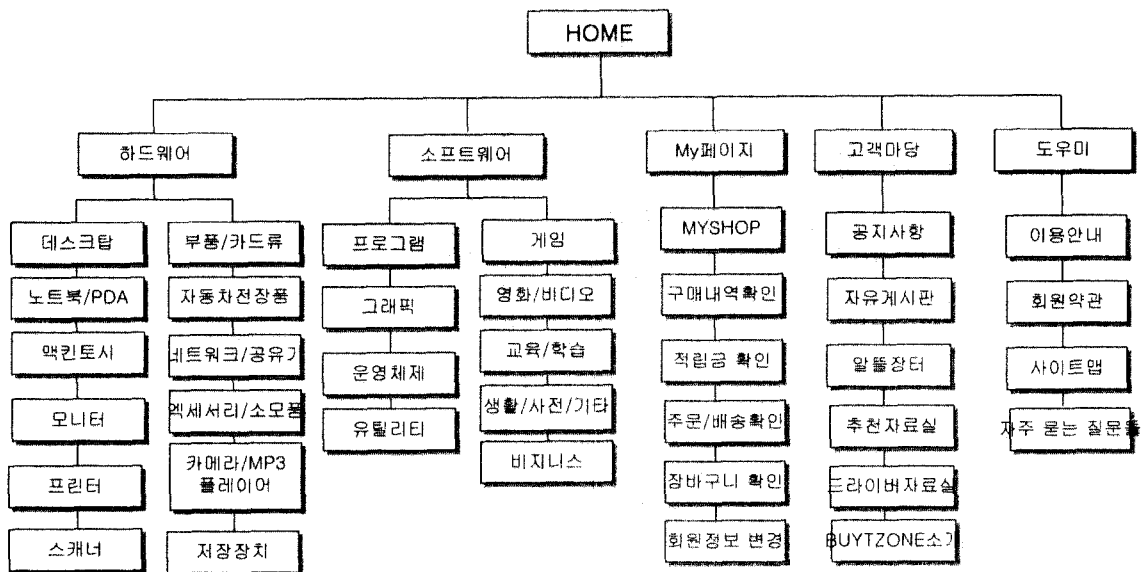
(1) 분석대상의 선정 및 사이트의 특성분석

본 연구의 분석대상은 온라인 쇼핑몰에서 컴퓨터와 주변기기를 전문적으로 판매하는 B쇼핑몰 사이트에서 1달 동안 접속하여 물건을 구매한 사용자

의 로그파일을 분석대상으로 선정하였다. 분석대상을 구체적으로 살펴보면 구매한 사람은 총 242명이며 이들이 B쇼핑몰 웹사이트에서 탐색한 로그파일은 총 13,000건이다. B 쇼핑몰의 구조를 한눈에 파악하기 위한 트리형태로 표현하면 <그림 3>과 같다.

<그림 3>에서 사이트의 구조를 살펴보면 크게 하드웨어와 소프트웨어를 판매하고 있으며 My페이지와 고객마당 도우미 등의 기능을 가지고 있다. 하드웨어와 소프트웨어의 경우 카테고리별로(데스크탑, 노트북, 맥킨토시 등등) 웹페이지를 작성하였으며 각 카테고리를 클릭하면 이를 구성하는 제품으로 연결되고 제품을 각 제품은 지불시스템과 연결되어 있다. My페이지의 경우 다음 "4.1.2 사이트의 특성 및 구매절차 분석"에서 자세하게 설명하겠지만 B쇼핑몰에서 고객들에게 개인화를 제공하는 특성중의 하나이다.

<그림 3> B 쇼핑몰의 사이트 구조



웹사이트를 효과적으로 분석하기 위해서는 B 쇼핑몰 사이트의 특성을 분석하는 것이 중요하다. B 쇼핑몰 사이트의 특성은 컴퓨터 및 컴퓨터 관련 주변기기를 판매, 경매의 쇼핑을 제공할 뿐만 아니라 원하는 제품검색, 공동구매, 채팅, 쇼핑, 고객상담 등을 동시에 이용할 수 있다. 따라서 물건을 구입하고자 하는 사용자들은 이 사이트에서 원하는 물건을 검색하고 이를 지불하는 구매의 모든 과정을 해결할 수 있다. 특히 B사이트에서는 "My페이지"라는 기능을 활용하여 고객들의 구매과정에서 다른 사이트와 차별화를 시도하고 있다. 세부적으로 살펴보면, 회원들이 쇼핑을 하다 가격변화가 심한 상품이나 비교가 필요한 상품들을 미리 등록해두고 필요시 곧바로 상품을 주문하고 결제할 수 있는 서비스를 제공하고 있다. 또한 한 번 장바구니에 담은 상품은 결제 또는 삭제하지 않는 한 계속 보관되므로 나중에 상품을 다시 검색하지 않고도 바로 구매하실 수 있다. 이에 본 연구의 목적이 웹로그 파일을 중심으로 고객의 구매 패턴을 분석하는 것이므로 B쇼핑몰 사이트에서 물건을 제시하는 사이트와 물건을 홍보하는 사이트를 전략적으로 활용할 핵심적인 콘텐츠를 중심으로 이들간의 연계성을 분석하고자 한다.

(2) 웹로그 자료의 전처리

분석대상 사이트를 선정하고 이들의 특성을 분석하였으면 다음으로는 웹로그 분석을 실시하기전에 전처리 과정이 필요하다. 왜냐하면 일반적인 웹로그의 경우 이미지 파일(gif, jpg 등), Include 파일(css, js 등), 자바 애플릿(cg), CGI 파일 등이 포함되어 있다. 따라서 웹로그는 그 자료의 수가 많으나 실질적으로 분석대상이 되는 정보를 가지고

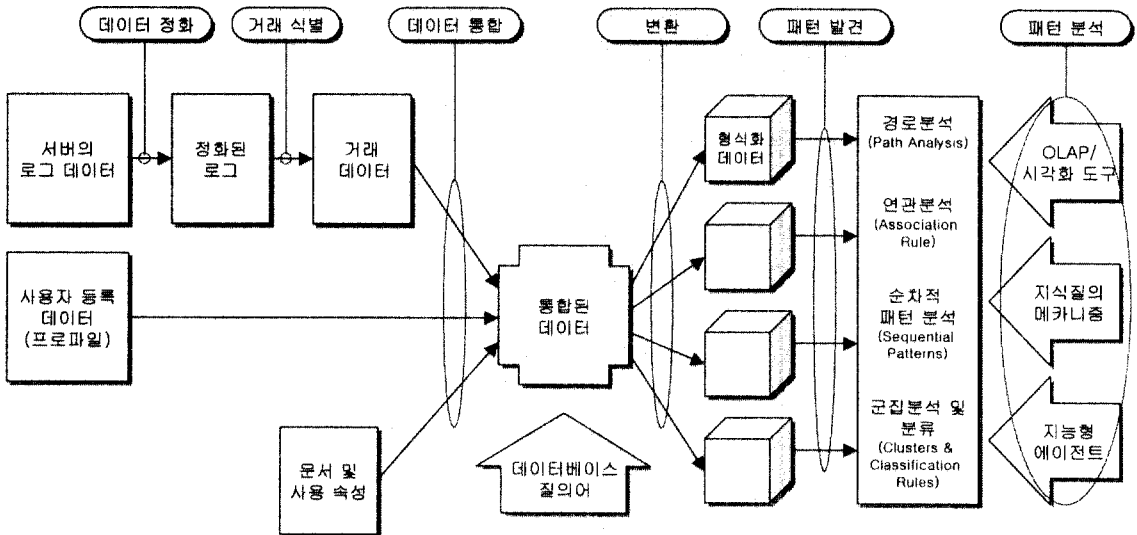
있는 레코드 수는 소수에 불과하다. 따라서 분석목적에 따라 로그 파일 중에서 적합한 항목을 선택하여 분석하여야 한다. 본 연구의 경우 사용자의 구매 패턴을 분석하여 연관성을 파악하는 것이 목적이기 때문에 사용자 IP 필드, 사용자가 클릭한 메뉴에 대한 필드 등이 분석에 필요하다. 따라서 트랜잭션에서 발생한 자료 중에서 등은 아무런 의미가 없기 때문에 삭제해야 한다. 이상의 과정을 거쳐서 관계형 데이터베이스 형태로 자료를 변환시켜 주어야 한다. 본 연구에서는 웹로그를 전처리하는 과정에서 불필요한 파일 등을 삭제하기 위해서는 "이지마이너(EasyMiner)"라는 프로그램을 활용하였다. 이지마이너는 자료의 전처리 과정을 담당하는 프로그램으로써 기존의 사이트에 접속하는 사람들의 로그파일을 전처리해서 일정한 관계형 데이터 테이블 형태로 보여주는 프로그램이다. 이에 본 연구에서는 이 프로그램을 활용하여 1차 전처리를 실시하였다.

3.2 인과관계 추출

실제 쇼핑몰의 웹로그를 대상으로 고객의 구매 패턴을 분석하고 구매에 영향을 미치는 주요한요인 및 이들의 인과관계를 도출하였다.

(1) 웹로그분석

일반적으로 웹로그 분석은 다양한 방법으로 이루어지나 본 연구에서는 Cooley(1997) 등이 제시한 일반적인 웹로그 분석 틀(Architecture)을 이용하기로 하였다. 다음 <그림 4>에는 이 같은 일반적인 웹로그 분석 틀이 제시되어 있다(Cooley, et al., 1997).

〈그림 4〉 웹로그 분석 기본 절차³⁾

〈그림 4〉를 구체적으로 살펴보자. 먼저 데이터베이스에 저장되어 있는 웹로그 파일에서 정보를 가져다주지 못하는 로그는 제거되고 의미 있는 로그만을 전처리하여 접속자별로 로그를 배열한다. 배열된 로그는 분석에 용이한 데이터로 전환하는데 이때 결함(Aggregation), 필터링(Filtering), 통합(Merge) 기법을 이용하게 된다. 분석하고자 하는 데이터를 확정하였으면 이를 활용하여 데이터베이스에서 패턴과 규칙을 도출하기 위하여 다양한 분석방법을 통하여 분석을 실시하게 된다. 이때 연관규칙분석, 순차적 패턴분석, 군집 및 분류 분석이 이용된다. 마지막으로 분석된 결과에서 의사결정에 도움이 되는 지식을 획득하게 되는 것이다. 본 연구에서도 이상의 과정을 이용하여 웹마이닝

분석을 실시하였다.²⁾

(2) 인과관계 도출

실제 웹쇼핑몰의 로그데이터를 분석하여 실제 고객들이 구매하는 과정에서 상대적으로 빈번히 활용하는 기능을 중심으로 이들 사이의 인과관계를 도출하기 위하여 연관규칙 분석을 실시하였다. 이때 사용한 알고리즘은 Apriori로 이에 대하여서는 앞의 연구배경에서 충분히 설명하였다(Agrawal and Srikant, 1994; Agrawal et al., 1993(a,b); Park et al., 1995; 박종수, 1996, 1998). 〈그림 5〉에는 분석결과가 제시되어 있다.

〈그림 5〉에서 제품상세설명→구매인 경우 지지도

2) 웹로그분석을 위하여 본 연구에서 사용되는 툴은 SPSS사에서 출시한 클레멘타인 5.2를 사용하였다. 클레멘타인은 데이터 마이닝에 용이한 기술을 통합한 패키지로써 현재 산업계에서 많이 활용하고 있는 프로그램으로 개체지향적(Object Oriented)이며 초보자가 사용하기 용이하다는 장점이 있다.

3) R. Cooley, et al., "Web Mining: Information and Pattern Discovery on the World Wide Web," ICTAI97, pp.26-32.

[illegible]

매는 0.8 정도의 인과관계로 설명할 수 있다. 따라서 이러한 과정을 통하여 분석된 결과를 중심으로 주요요인과 이들의 인과관계 값을 정리하면 다음에 제시되는 <표 3>과 같다.

주요요인	인과관계값
이미지확대기능 → 구매	0.9
가격비교기능 → 구매	0.8
제품상세설명 → 구매	0.8
장바구니기능의 활용 → 거래승인	0.7
장바구니기능의 활용 → 결제수단의 다양성	0.7
등록 및 인증 → 개인화기능	0.9
개인화기능 → 거래승인	0.9
이벤트 및 경품기능 → 등록 및 인증	0.8
이벤트 및 경품기능 → 파법활용	1.0
이벤트 및 경품기능 → 메인상품전시	0.6
메인 홈페이지 → 등록 및 인증	0.5
메인 홈페이지 → 메인상품정리	0.5
메인상품정리 → 제품분류기능	0.7
제품분류기능 → 이미지 확대 기능	0.8

3.3 퍼지인식도의 작성

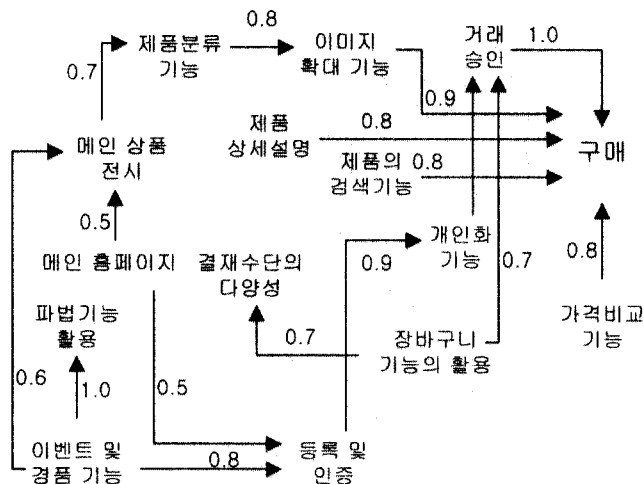
퍼지인식도를 작성하는 방법에는 의사결정자의 주관적인 판단에 기초하거나(Axelrod, 1976; Eden et al., 1979), 설문서를 작성하여 그 결과를 종합함으로써 결정하거나(Montazemi & Conrath, 1986) 또는 인공신경망 등의 다양한 데이터마이닝 방법을 적용하는 방법(Caudill, 1990) 등이 있다. 본 연구에서는 제안하는 퍼지인식도의 작성은 크게 세 과정을 통하여 이루어진다. 첫째, 핵심(Core) 퍼지인식도의 작성으로 이는 웹로그 분석을 통하여 도출된 결과를 중심으로 작성된다. 둘째 기존문헌을 분석하여 웹사이트를 구성하는 중요요인을 도출하고 이를 중심으로 전문가들에게 문의하여 슈퍼(Super) 퍼지인식도를 작성한다. 셋째, 핵심 퍼지인식도와 슈퍼 퍼지인식도를 합하여 웹쇼핑몰에서 사용자의 구매패턴을 나타내는 통합 퍼지인식도를 도출한다.

(1) 핵심 퍼지인식도의 작성

본 연구에서는 기존의 퍼지인식도의 연구와는 다르게 실제 웹로그 자료를 분석하여 주요 요인을 도출하였으며, 요인들간의 연관관계를 추출하는 Apriori 알고리즘을 이용하여 이들 요인들간의 인과관계를 추출하였는데 그 결과가 <표 3>에 요약되어 있다. <표 3>을 바탕으로 작성된 퍼지인식도에는 웹쇼핑몰에서 실제사용자들의 실제적인 구매패턴이 반영되어 있다는 장점이 있다. 이때의 퍼지인식도를 본 연구에서는 핵심 퍼지인식도라고 정의하였으며 이를 도시하면 <그림 6>과 같다.

<그림 6>을 구체적으로 살펴보자. 메인 홈페이지의 변화는 메인상품 전시와 제품의 분류기능 및 이미지 확대기능에 영향을 미치게 되고, 이벤트 및 경품기능은 메인상품 전시와 등록 및 인증, 팝업(pop-up) 기능의 활용에 영향을 주어 전체적으로 구매에 영향을 미치게 된다. 그리고 개인화 기능이나 장바구니 기능의 활용은 거래승인에 영향을 주

<그림 6> 웹쇼핑몰의 사용자 구매에 대한 핵심 퍼지인식도



어 구매에 영향을 미침을 알 수 있다. 또한 제품에 대한 상세 설명 및 제품의 검색기능도 구매에 영향을 미치는 것으로 분석되었다.

(2) 수퍼 퍼지인식도의 작성

웹로그 분석을 통하여 추출한 인과관계를 중심으로 작성한 핵심 퍼지인식도의 경우 사용자들의 실제 웹사이트 방문거래를 바탕으로 하였으므로 객관적이며 실제자료에 근거한 형태지식을 표현할 수 있다. 하지만 웹쇼핑몰 구축에 필요한 전문가의 지식은 반영되지 않는다는 한계점을 가지고 있다. 따라서 본 연구에서는 웹쇼핑몰에서 사용자의 구매에 대한 긍정적인 요인을 반영하기 위하여 기존문헌을 통하여 웹 쇼핑몰 구축에 대한 주요요인을 도출하고 이를 중심으로 전문가들의 의견을 반영하여 퍼지인식도를 작성하였는데 이를 수퍼 퍼지인식도라고 명하였다.

먼저, 웹 쇼핑몰 사이트를 구축에 대한 연구로써 Lohse & Spiller(1998)를 필두로 하여 많은 전문가들이 실증자료를 바탕으로 구체적인 방법론을 제시하고 있다(김진우 & 문재윤, 1999; 김광용 & 김기수, 1999; 안준모 & 한상록, 2000; Forbes & Rothschild, 2000; Huizingh, 2000; Jarvenpaa & Todd, 1997; Liang and Lai, 2000; Spiller & Lohse, 1998; Wells & Fuerst, 2000; Zhang et al., 2000). 이에 본 연구에서는 이들의 연구를 근간으로하여 웹 사이트의 영향요인을 파악하고 이들 요인이 사용자의 만족도 측면과 갖는 인과관계를 바탕으로 퍼지인식도를 작성하고자 한다. 사용자의 만족도는 이미 많은 연구에서 소비자의 구매행동을 촉발시킬 수 있는 중요한 요인으로 삼고 있기 때문에 본 연구에서 채택한 것이다.

먼저, 상품측면에서 보자. Spiller & Lohse(1998)

는 제품의 개수가 많은 큰 쇼핑몰의 경우에는 소비자가 제품을 찾기가 어렵기 때문에 메인 화면에서의 상품의 정리가 중요하다고 주장하였다. 따라서 다음과 같은 인과관계를 생각할 수 있다.

제품의 개수 $\xrightarrow{+}$ 하이퍼링크의 수

Huizingh(2000)은 웹 사이트의 내용과 디자인의 형태가 해당사이트의 크기와 매우 밀접하다고 강조하고 있기 때문에 이러한 인과관계는 매우 타당하게 보인다. Jarvenpaa & Todd(1997)은 그들의 연구에서 제품에 대한 인지는 사용자의 만족도에 많은 영향을 미친다고 보고하고 있다. 따라서, 다음과 같은 인과관계를 고려할 수 있다.

하이퍼링크를 이용한 제품설명 $\xrightarrow{+}$ 콘텐츠의 풍부성

특히 하이퍼 링크는 새롭고 유용한 정보를 얻도록 사용자를 돕는 기능을 한다. De Troyer & Leune(1998)은 기존의 웹 사이트가 사용자 중심이 아닌 웹을 만드는 사람의 입장에서 구축되어 많은 문제점을 가지고 있지만 적절한 하이퍼링크를 통해 사용자 관점에서의 편리한 웹 디자인을 가능하게 한다고 하였다. 이와 같이 일반적으로 하이퍼링크는 유사한 내용 또는 관련내용을 포함하고 있기 때문에 사용자가 특정한 제품과 관련된 정보를 찾는 노력을 감소시킨다. 따라서 다음과 같은 인과관계가 형성된다.

하이퍼링크 $\xrightarrow{-}$ 탐색노력

Larson & Czerwinski(1998)는 웹 사이트 메뉴 디자인에서 넓이와 깊이가 다른 웹 사이트를 디자인 하여 이러한 인과관계가 실증적으로 존재함을

증명하였다. 또한, 웹 사이트에서 그림을 사용하면 천 마디 말보다 효과가 우수하다. 그러나, 너무 작은 그림은 질이 낮아 전달하려는 메시지가 모호해지며, 큰 그림은 다운로드 받는데 너무 시간이 오래 소요된다. 따라서, 다음과 같은 인과관계가 형성된다(김진우 & 문재윤, 1997).

멀티미디어기능의 사용 ----> 사이트 반응속도

배너광고나 일반광고 측면에서 볼 때 어떤 사이트는 자바, 사운드, 애니메이션 등과 같은 것들을 적절히 다루지 못함으로써 사이트의 속도가 저하되어 사용자들을 짜증나게 한다. 이것은 그림의 크기가 크면 사이트의 반응속도가 감소하는 인과관계와 연관지어서 생각해 볼 수 있다. 따라서, 다음과 같은 인과관계를 생각할 수 있다.

오디오, 애니메이션, 그래픽, 비디오 사용 ----> 사이트 반응속도

Zhang et al.(2000) 역시 사용자의 만족과 불만족을 나누는 요인으로 탐색활동에 있어서의 신속

성을 주장하였다. 따라서, 다음과 같은 인과관계 도출이 가능하다.

사이트 반응속도 ----> 사용자 만족도

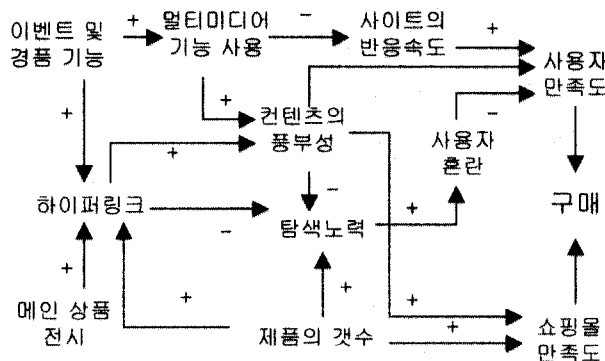
한편, 모든 사용자가 웹 사이트의 첫 화면으로 들어오는 것은 아니다. 검색 엔진을 사용하면 웹 사이트의 가장 끝 정보로 들어올 수도 있다. 이때 네비게이션 버튼이 없다면 사용자는 더 이상의 정보검색이 불가능하다. 따라서, 검색 기능이 있는 웹 사이트에서 웹 페이지의 일관성을 유지한다는 것은 상당히 어렵지만 이것이 이루어질 경우 사용자들은 매우 편리하게 서핑이 가능하다.(Lohse & Spiller, 1998).

제품 검색 기능 ----> 사용자의 만족도

이상과 같은 웹 사이트의 영향요인들의 인과관계와 웹마이닝 기법을 적용한 주요요인 도출을 중심으로 퍼지인식도를 작성하고자한다.

이상의 내용을 바탕으로 웹쇼핑몰에서 사용자의 구매패턴에 대한 수퍼퍼지인식도를 도출하면 <그림 7>

<그림 7> 웹쇼핑몰에서 사용자 구매패턴에 대한 수퍼 퍼지인식도



과 같다.

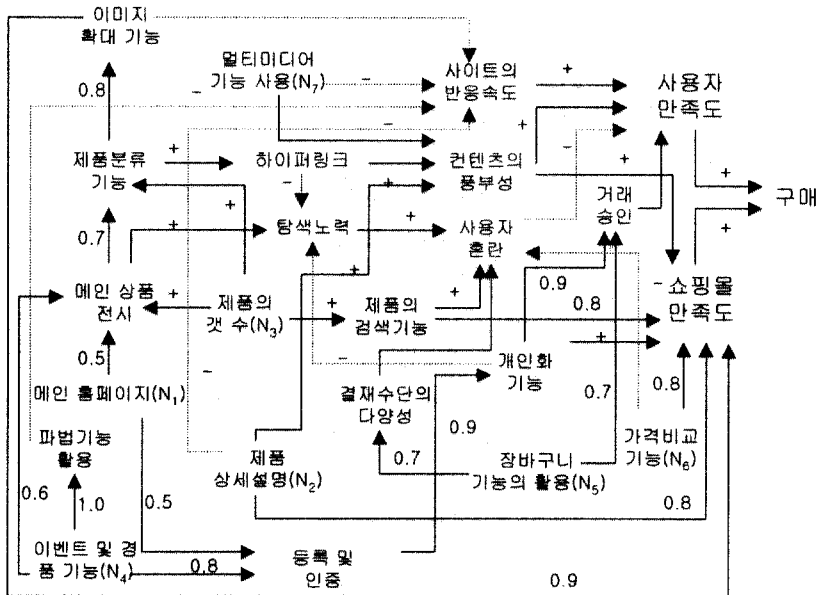
〈그림 7〉을 살펴보자. 먼저, 제품의 개수가 많은 큰 쇼핑몰의 경우에는 소비자가 제품을 찾기가 어렵기 때문에 메인 화면에서의 상품의 정리 기능은 하이퍼링크의 수를 많게 만들어 양(+)의 영향을 미치지만 사용자가 원하는 물건에 대한 탐색노력을 줄일 수 있으므로 음(-)의 영향을 미치게 된다. 또한 이벤트 및 경품기능을 사용은 멀티미디어 기능의 사용에 양(+)의 영향을 미치게 되지만 이는 사이트의 반응속도에 음(-)의 영향을 미치게 된다. 그리고 콘텐츠의 풍부성은 전체적으로 쇼핑몰의 만족도에는 양(+)의 영향을 가져다주나 탐색노력에는 음(-)의

영향을 미치게 되고 이는 사용자의 혼란을 가져와 사용자 만족도에는 음(-)의 영향을 미치게 된다.

(3) 통합 퍼지인식도의 작성

웹쇼핑몰 사용자의 구매패턴을 제시하는 핵심 퍼지인식도의 경우 객관적인 요인을 반영할 수 있다는 장점이 있지만 웹쇼핑몰에서 사용자들의 주관적인 요인을 반영할 수 없다는 한계점을 가지고 있다. 따라서 이 두 퍼지인식도를 합하여 웹쇼핑몰에서 사용자의 구매패턴을 나타내는 통합 퍼지인식도를 작성할 수 있는데 〈그림 8〉과 같다. 〈그림 8〉

〈그림 8〉 웹 쇼핑몰 사용자 구매패턴에 대한 통합 퍼지인식도⁴⁾



4) 이때 '+' 혹은 '-'로 표현된 부분은 수퍼 퍼지인식도에서 가져온 것으로 이들간의 인과관계값의 크기는 웹로그 분석을 통하여 도출할 수 없다. 따라서 이들간의 인과관계가 비례관계에 있으면(+), 반비례관계에 있으면(-)로 표현하였으며, 〈표 4〉에서는 '+' = 1로 '-' = -1로 인과관계값을 설정하였다.

에 제시되어 있는 통합 퍼지인식도의 경우 앞의 핵심 퍼지인식도와 수퍼 퍼지인식도 보다 그 형태는 복잡하지만 웹쇼핑몰에서 사용자의 구매패턴에 대한 지식을 잘 표현하고 있다. 특히 사용자의 구매에 영향을 주는 두 요인으로 쇼핑몰 만족도와 사용자 만족도를 표현함으로써 이들간의 연계성을 표현하였다.

IV. 시뮬레이션 및 전략제시

본 장에서는 <그림 8>에 제시되어 있는 퍼지인식도를 바탕으로 다양한 조건에서 시뮬레이션을 실시하고 이 결과를 바탕으로 웹쇼핑몰 사용자의 구매에 대한 전략을 제시하고자 한다. 이때 퍼지인식도를 추론하는 방식으로는 계층화된 퍼지인식도를 이용하고자 한다. 이는 기존의 퍼지인식도 추론과정

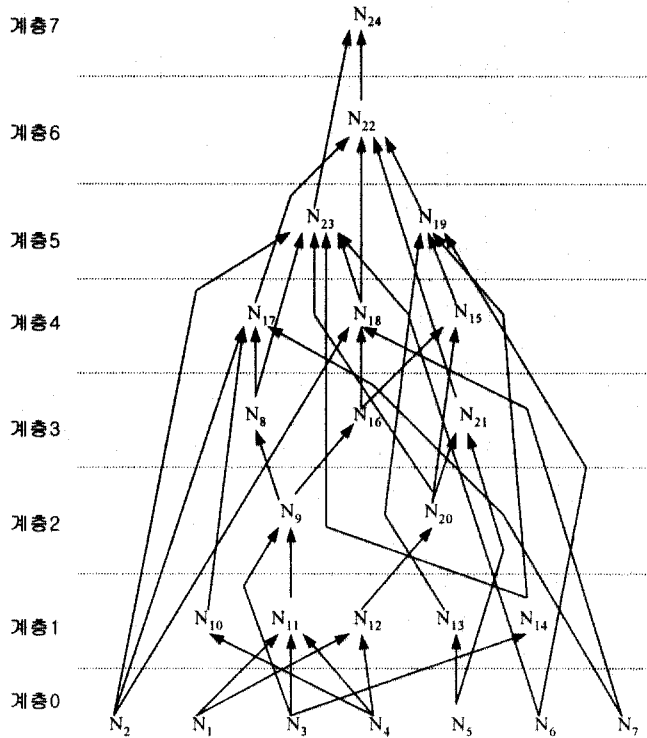
은 추론시 단순한 임계치 방식을 사용하기 때문에 추론의 경직성을 초래하는데 이는 너무 단순한 추론결과를 도출하여 현실 상에서 의미있는 결과를 얻기가 어렵다. 여기에 관한 자세한 내용을 알기 위해서는 Zhang & Chen(1988), Kosko(1986, 1992), Taber(1991) 등의 연구를 참조하기 바란다. 계층화된 추론 알고리즘은 이진창 & 조형래(1998)가 개발한 것으로 기존의 퍼지인식도 추론 방식이 가지는 동기화의 문제 및 추론결과의 단순성의 문제를 해결한 알고리즘이다. 다음 <표 4>에는 <그림 8>에 제시된 퍼지인식도에 대한 연관행렬을 표현하고 있다.

먼저, 계층화된 추론을 하기 위하여 주어진 연관행렬을 바탕으로 각 노드들에 대한 계층을 구하는 과정을 정리하면 계층 0은 노드 $N_1, N_2, N_3, N_4, N_5, N_6, N_7$, 계층 1은 $N_{10}, N_{11}, N_{12}, N_{13}, N_{14}$, 계층 2는 N_9, N_{20} , 계층 3은 N_8, N_{16}, N_{21} 은 계층 2, 계층 4는 N_{15}, N_{17}, N_{18} , 계층 5는 N_{19} .

<표 4> 사례 퍼지인식도에 대한 연관행렬

	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20	N21	N22	N23	N24
N1	0	0	0	0	0	0	0	0	0	0	0.5	0.5	0	0	0	0	0	0	0	0	0	0	0	0
N2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	1	0	0	0	0	0	0.8	0
N3	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
N4	0	0	0	0	0	0	0	0	0	1	0.6	0.8	0	0	0	0	0	0	0	0	0	0	0	0
N5	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0	0.7	0	0	0
N6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0.8	0
N7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	1	0	0	0	0	0	0	0
N8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0.9	0
N9	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
N10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
N11	0	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.9	0	0	0	0	0
N13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
N14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
N15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
N16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	1	0	0	0	0	0	0	0
N17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
N18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
N19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0
N20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0.9	0	1	0
N21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
N22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
N23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
N24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

〈그림 9〉 계층화된 퍼지인식도



N_{23} , 계층 6은 N_{22} , 계층 7은 N_{24} 에 속함을 알 수 있다. 이상의 과정을 통하여 도출한 각 노드의 값을 계층을 중심으로 퍼지인식도로 작성하면 〈그림 9〉와 같다.

〈그림 9〉에서 입력노드가 속하는 계층은 0임을 알 수 있다. 또한 결과노드가 속하는 계층은 해당 노드에 대한 원인노드가 속하는 계층보다 상위 계층에 속함을 알 수 있다. 이와 같이 계층화 시킨 다음에는 하이퍼탄젠트(hyper-tangent) 함수를 바탕으로 추론을 한다(이건창 & 조형래, 1998). 이와 같이 구축된 계층화된 퍼지인식도는 입력값인 $N_1 \sim N_7$ 의 영향력의 크기를 조정함으로써 최종적으로 만족도 노드 N_{24} 의 크기를 변화시킬 수 있다.

본 연구에서는 각각의 입력값의 범위를 0.0~1.0사이에서 변화시키면서 시뮬레이션을 실시하였다. 구체적으로 시뮬레이션을 실시한 결과를 살펴보자.

① 시나리오 1-메인홈페이지의 변화

모든 고객들은 메인 홈페이지에 접속한 뒤 일정 동안 항해를 하고 구입하기를 원하는 물건을 결정하게 된다. 따라서 메인홈페이지가 쇼핑몰에서 매우 중요한 역할을 담당하게 된다. 특히 마케팅 전략의 하나로써 쇼핑몰을 방문하는 고객에게 가격의 상대적으로 저렴한 물건을 초기화면에 팝업(Pop-up)

〈표 5〉 시뮬레이션 입력값과 결과값

입력값		결과값	
변수기호	설명	변수기호	설명
N ₁	메인 홈페이지	N ₂₄	구매
N ₂	제품에 대한 상세한 설명 기능		
N ₃	제품의 개수		
N ₄	이벤트 및 경품기능		
N ₅	장바구니 기능의 활용 정도		
N ₆	가격의 비교기능		
N ₇	멀티미디어(애니메이션, 자바)기능		

기능을 함께 활용하여 제시함으로 고객들의 구매욕구를 유발할 수 있으며, 상품을 전시해두는 메인상품 전시 페이지로 고객의 향해를 유도하는 전략이 가능하다. 이에 메인홈페이지를 변화시켰을 때 미치는 영향관계를 살펴보기 위하여, 다른 조건은 일반적인 보통값인 0.5로 두고 메인 홈페이지 변화값은 1로 두고 시뮬레이션을 실시하였다. 이때 입력값은 아래와 같다.

$$N_1=1.0, N_2=0.5, N_3=0.5, N_4=0.5, \\ N_5=0.5, N_6=0.5, N_7=0.5$$

이를 바탕으로 추론 시뮬레이션을 실시하면 〈그림 10〉과 같은 결과를 얻을 수 있다.

메인 홈페이지를 변화시켰을 때 사용자들의 탐색 노력에 영향을 미치게되고 또한 판매하는 상품을 분류하는 요인에도 영향을 미쳐서 최종적으로 구매에 0.9303 만큼의 변화를 미치는 것으로 분석되었다.

② 시나리오 2-웹 쇼핑물에서의 판매 제품의 개수의 변화

웹 쇼핑물에서 가장 빈번이 일어나는 일 중의 하나가 판매하는 제품의 개수를 변화시키는 일이다. 특히 새로운 인기 상품이 등장하거나 더 이상 판매가 되지 않는 제품이 있는 경우 이러한 제품은 웹 쇼핑물에서 제외된다. 이러한 경우에 고객의 구매에 어떠한 영향을 미치는지를 분석하기 위하여 웹 쇼핑물에 전시되는 제품의 개수를 변화시켜 보았

〈그림 10〉 메인 홈페이지를 변화시켰을 때

추론	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20	N21	N22	N23	N24
계층 0	1.000	0.500	0.500	0.500	0.500	0.500	0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
계층 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.462	0.874	0.716	0.336	0.462	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
계층 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.805	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.568	0.000	0.000	0.000	0.000
계층 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.567	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.667	0.000	0.000	0.000	0.000	0.336	0.000	0.000	0.000
계층 4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.844	0.000	-0.966	0.931	0.000	0.000	0.000	0.000	0.000	0.000
계층 5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.497	0.000	0.000	0.000	0.997	0.000
계층 6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.663	0.000	0.000
계층 7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.930
최종	1.000	0.500	0.500	0.500	0.500	0.500	0.500	0.567	0.805	0.462	0.874	0.716	0.336	0.462	-0.844	0.667	-0.966	0.931	-0.497	0.568	0.336	0.663	0.997	0.9303

다. 이때의 입력값은 아래와 같다.

기능의 사용

$$N_1=0.5, N_2=0.5, N_3=1.0, N_4=0.5, \\ N_5=0.5, N_6=0.5, N_7=0.5$$

이를 바탕으로 추론 시뮬레이션을 실시하면 <그림 11>과 같은 결과를 얻을 수 있다.

분석결과를 살펴보면 새로운 제품이 전시됨으로 이에 대한 상세설명기능, 제품의 검색 기능, 제품의 분류기능이 변화되고 최종적으로 0.9455으로 앞의 메인 홈페이지를 변화시켰을 때보다 최종적인 구매 값이 조금 더 향상된 것을 알 수 있다. 이는 제품의 변화에 소비자들의 구매에 민감하게 영향을 미친다는 뜻이다. 특히 새로운 제품이 쇼핑몰에 제시되면 이 제품에 대한 구매는 향상될 것이다.

소비자들의 의식수준이 향상됨에 따라 제품을 구매할 때 이에 대한 상세한 설명을 원하는 고객들이 증가하고 있다. 따라서 쇼핑몰에서도 이러한 소비자들의 요구를 반영하기 위하여 제품에 대한 상세 설명을 추가하여 웹 쇼핑몰 전체적으로 콘텐츠의 양을 향상시킬 필요가 있다. 또한 제품에 대한 사실적인 묘사를 위하여 제품에 대한 전시기능을 향상시키기 위하여 시각적으로 해상도가 높은 이미지를 사용하거나, 음향 효과 등 다양한 기능을 사용하기도 한다. 따라서 제품에 대한 상세설명과 멀티미디어 효과 등을 측정하기 위하여 시뮬레이션을 실시하기 위하여 아래의 값을 입력하였다.

$$N_1=0.5, N_2=1.0, N_3=0.5, N_4=0.5, \\ N_5=0.5, N_6=0.5, N_7=1.0$$

③ 시나리오 3-제품의 상세설명 및 멀티미디어

<그림 11> 전시하는 제품의 수를 변화시켰을 때

추론	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20	N21	N22	N23	N24
계층 0	0.500	0.500	1.000	0.500	0.500	0.500	0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
계층 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.462	0.914	0.572	0.336	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
계층 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.927	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.473	0.000	0.000	0.000	0.000	0.000
계층 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.630	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.729	0.000	0.000	0.000	0.336	0.000	0.000	0.000	0.000
계층 4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.835	0.000	-0.926	0.939	0.000	0.000	0.000	0.000	0.000	0.000
계층 5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.761	0.000	0.000	0.000	0.983	0.000	0.000
계층 6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.804	0.000	0.000	0.000
계층 7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.945	0.000
최종	0.500	0.500	1.000	0.500	0.500	0.500	0.500	0.630	0.927	0.462	0.914	0.572	0.336	0.000	-0.835	0.729	-0.926	0.939	-0.761	0.473	0.336	0.804	0.983	0.945

<그림 12> 제품의 상세설명 기능 및 멀티미디어 기능을 변화하였을 때

추론	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20	N21	N22	N23	N24
계층 0	0.500	1.000	0.500	0.500	0.500	0.500	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
계층 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.462	0.782	0.572	0.336	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
계층 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.781	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.473	0.000	0.000	0.000	0.000	0.000
계층 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.554	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.653	0.000	0.000	0.000	0.000	0.336	0.000	0.000	0.000
계층 4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.810	0.000	-0.988	0.980	0.000	0.000	0.000	0.000	0.000	0.000
계층 5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.750	0.000	0.000	0.000	0.982	0.000	0.000
계층 6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.796	0.000	0.000	0.000
계층 7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.945	0.000
최종	0.500	1.000	0.500	0.500	0.500	0.500	1.000	0.554	0.781	0.462	0.782	0.572	0.336	0.000	-0.810	0.653	-0.988	0.980	-0.750	0.473	0.336	0.796	0.982	0.945

분석결과 최종적인 구매정도값은 0.9445로써 최초의 단순히 메인 홈페이지를 변화하였을 때보다는 많은 변화가 일어난 것을 알 수 있다. 하지만 앞서 제품의 갯수를 변화시켰을 때 최종적인 구매정도값인 0.945보다는 약간 적은 것으로 분석되었다. 이러한 이유로는 제품에 대한 상세설명과 제품의 사실성을 위한 멀티미디어의 사용은 제품에 대한 사실성을 높이기 때문에 제품에 대한 만족도는 높힐 수 있으나 사이트의 반응속도가 느려지고 하이퍼링크가 많아짐으로 사용자 만족도는 감소하기 때문이다. 따라서 실제 웹쇼핑몰에서는 판매하는 제품에 대한 사실감과 사이트의 반응속도를 적절히 배분하여야 한다.

④ 시나리오 4-장바구니 기능 및 가격 비교 기능의 변화

웹쇼핑몰의 가장 큰 장점중의 하나가 제품을 한 곳에서 손쉽게 구매할 수 있다는 것이다. 따라서 기존의 전통적이 상거래에서 발생하는 쇼핑의 개념이 웹쇼핑몰에서도 가능한데, 웹쇼핑몰에서는 구매하고자하는 제품을 다양한 쇼핑물을 돌아다니면서 비교할 수 있다. 따라서 소비자들은 자신이 구입하기 원하는 제품을 서로 가격을 비교하기도하기도 한다.

또한 대부분의 쇼핑물에서 고객들이 구입하기 원

하는 제품에 대하여 마지막 계산까지의 과정을 편리하게 지원하기 위하여 장바구니 기능을 가지고 있다. 따라서 이러한 기능들이 변화하였을 때 구매에 미치는 영향을 분석하기 위하여 아래의 값을 입력하였다.

$$N_1=0.5, N_2=0.5, N_3=0.5, N_4=0.5, \\ N_5=1.0, N_6=1.0, N_7=0.5$$

분석결과 가격비교기능과 장바구니 기능을 활성화할 때 최종적인 구매값은 0.954로 높게 제시되었다. 이는 가격비교기능으로 제품의 만족도가 향상되었으며, 장바구니의 활용으로 사용자의 만족도가 향상되었기 때문이다.

V. 결론 및 향후 연구방향

기존의 웹 쇼핑물에 대한 연구들은 대부분, 웹 사이트 구축을 위한 영향요인이 무엇인지를 탐색적으로 찾는 분야에 집중되어 있다. 또한 일부 연구들은 제품의 다양성, 서비스, 촉진, 항해량, 편리성, 사용자 인터페이스 등의 디자인 측면에서 웹 쇼핑물의 중요성을 강조하였지만, 문제는 그러한

〈그림 13〉 장바구니의 기능 및 가격 변화기능을 변화하였을 때

수준	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20	N21	N22	N23	N24
계층 0	0.500	0.500	0.500	0.500	1.000	1.000	0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
계층 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.462	0.782	0.572	0.604	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
계층 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.781	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.473	0.000	0.000	0.000	0.000
계층 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.564	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.653	0.000	0.000	0.000	0.000	0.604	0.000	0.000	0.000
계층 4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.810	0.000	-0.914	0.929	0.000	0.000	0.000	0.000	0.000	0.000
계층 5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.835	0.000	0.000	0.000	0.965	0.000
계층 6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.897	0.000	0.000
계층 7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.954
최종	0.500	0.500	0.500	0.500	1.000	1.000	0.500	0.564	0.781	0.462	0.782	0.572	0.604	0.000	-0.810	0.653	-0.914	0.929	-0.835	0.473	0.604	0.897	0.965	0.954

연구결과를 실제로 웹쇼핑몰에 활용하기란 쉽지가 않다. 그 이유는 이들 요인들은 서로 복잡한 인과 관계를 가지고 있기 때문이다. 예를 들어, 사용자 인터페이스가 복잡하면 항해량이 늘어나 편리성이 감소하고, 제품이 늘어나더라도 검색엔진을 사용하면 상대적으로 항해량이 감소하게 되어 편리성이 증가하게 된다. 이와 같이 웹 쇼핑몰의 구성요소간에는 서로간에 복잡한 인과관계가 형성되어 있어서, 웹 쇼핑몰을 구축하여 소비자들의 만족도를 극대화하기 위해서는 구성요소간의 복잡한 인과관계를 면밀히 분석할 필요가 있다.

이에 본 연구에서는 퍼지인식도를 이용하여 보다 구체적이고 계량화된 웹 쇼핑몰 구축 방법을 제안하였다. 기존의 퍼지인식도에 대한 많은 연구들이 실제 데이터를 반영하지 못하고 기존연구나 연구자의 주관에 의하여 작성되는 한계를 가지고 있었지만, 본 연구에서는 이러한 한계점을 극복하기 위하여 실제 웹 쇼핑몰의 로그파일을 분석하여 이들의 중요속성을 도출하고 이를 데이터마이닝 기법을 적용하여 인과관계를 분석하고, 그 결과를 토대로 하여 퍼지인식도를 작성하였다. 본 연구의 공헌점은 다음과 같다.

첫째, 연구방법론상의 가치이다. 현재 국내외에서 이뤄지고 있는 웹 쇼핑몰에 대한 연구는 대부분이 성공적인 웹 사이트 구현을 위한 영향요인에 관한 연구이며 그 연구방법론은 설문조사를 통한 검증에 의존하고 있다. 그러나, 본 연구에서는 웹 쇼핑몰을 하나의 인과관계 시스템으로 간주하고 구성요소간의 인과관계를 퍼지인식도로 간주하여 최적의 산출물을 획득하기 위한 입력물의 조합을 제시할 수 있다.

둘째, 기존의 퍼지인식도의 연구 분야를 넓혔다는 사실이다. 본 연구는 최근 급부상하고 있는 웹

로그 파일을 대상으로 실제 데이터를 중심으로 퍼지인식도를 새롭게 접근했다는 면에서 퍼지인식도로 해결가능한 문제영역을 한단계 확장하는데 기여하였다. 특히 실제데이터에 근거한 인과관계의 분석과 더불어 기존문헌의 연구를 통하여 도출한 인과관계를 병합함으로써 실제 자료의 정성적인 요인도 반영하였다는 데에서 의미를 찾을 수 있다.

셋째, 웹쇼핑몰을 실제적으로 설계하는 실무적인 가이드 라인을 제공하였다. 특히 웹쇼핑몰의 특정 구성요소를 변화를 주었을 때에 그 결과가 최종적으로 어떻게 목적변수에 영향을 주는지를 확인할 수가 있다.

참 고 문 헌

- 김광용, 김기수, "인터넷 설문조사를 활용한 사이버 쇼핑물 디자인에 관한 연구," *경영정보학 연구*, 9(2), pp. 133-150.
- 김진우, 문재현, "시각 디자인 요소와 감성 요소간의 상관관계에 대한 연구," *경영정보학 연구*, 7(1), 1997, pp. 113-138.
- 박종수, 유원경, 홍기형, (1998), "연관규칙 탐사와 그 응용," *정보과학회지*, 16(9), pp. 37-44.
- 박종수(1996), "대용량 데이터베이스상의 효과적인 연관규칙 탐사를 위한 전지방법," *한국정보과학회 데이터베이스 연구회지*, 12(4), pp. 59-75.
- 안준모, 한상록(2000), "인터넷 쇼핑물 성공전략: 구성 디자인 특성과 마케팅 기능," *Information Systems Review*, 1(2), pp.97-105.
- 이건창, 조형래(1998), "계층화된 퍼지인식도를 이용한 추론 메카니즘에 관한 연구," *한국경영과학회*, 23(4), pp. 203-212.
- 한상린(1998), "Flow 개념을 이용한 인터넷 환경에서의 소

비자 구매의도 결정요인 분석," **한국마케팅학회 추계학술대회**.

- Agrawal, R. and R. Srikant(1994-a), "Fast Algorithms for Mining Association Rules in Large Databases," *Proceedings of the 20th International Conference on Very Large Databases*, pp. 115-121.
- Agrawal, R. and R. Srikant(1994-b), "Mining Sequential Patterns," Research report RJ 910, IBM Almaden Research Center.
- Agrawal, R., Imielinski, T. and A. Swami(1993-a), "Database Mining: A Performance Perspective," *IEEE Transactions on Knowledge and Data Engineering*, vol. 5, no. 6, pp. 914-925.
- Agrawal, R., Imielinski T., and A. Swami(1993-b), "Mining Association Rules between Sets of Items in Large Databases," *In Proc. of the ACM SIGMOD Conference on Management of Data*, pp. 207-216, Washington, D.C.
- Axelrod, R., *Structure of Decision: The Cognitive Maps of Political Elites*. Princeton, NJ: Princeton University Press, 1976.
- Banini, G.A. and R.A. Bearman(1998), "Application of Fuzzy Cognitive Maps to Factors Affecting Slurry Theology," *International Journal of Mineral Processing*, vol. 52, pp. 233-244.
- Borges, J. and M. Levene.(1999), "Mining Navigation Pattern Discovery from Internet Data," IBM, *Web Design Guideline*, http://www-3.ibm.com/ibm/easy/eou_ext.nsf/publish/572.
- Caudill, M.(1990), "Using Neural Nets: Fuzzy Cognitive Maps," *AI Expert*, June, pp. 49-53.
- Charu, C. Aggarwal and P.S. Yu(1998), "Mining Large Item sets for Association Rules. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, pp. 23-31.
- Chen, M.S., Han J., and P.S. Yu(1996), "Data Mining: An Overview from a Database Perspective," *IEEE Transactions on Knowledge and Data Engineering*, vol. 8, no. 6, pp. 866-883.
- Chen, M.S., Park, J.S. and P.S. Yu(1998), "Efficient Data Mining for Path Traversal Patterns," *IEEE Transactions on Knowledge and Data Engineering*, vol. 10, no. 2, pp. 209-221.
- Cooley, et al.(1997), "Web Mining: Information and Pattern Discovery on the World Wide Web," *ICTAI97*.
- Cooley, R., Tan, P.N. and J. Srivastava(1999), "Discovery of Interesting Usage Patterns from Web Data," Technical Report TR 99-022. University of Minnesota.
- Csikszentmihalyi, M.(1990), *Flow: The Psychology of Optimal Experience*, New York, Harper & Row Publisher.
- De Troyer, O.M.F. and C.J., Leune(1998), "WSDM: A User Centered Design Method for Web Site," *Computer Networks and ISDN Systems*, vol. 30, pp. 85-94.
- Edden, C., Jones, C., and Sims, D., *Thinking in Organizations*. MacMillan Press Ltd, London, England, 1979.
- Forbes, M.W. and M.L. Rothschild(2000), "Toward an Understanding of Consumer Experience on the Internet: Implications for Website Design," *Proceedings of the 33rd Hawaii International Conference on System Sciences: HICSS '2000*, pp. 1-10.
- Fu, Y., Sandhu, K. and M.Y. Shih.(1999), "Clustering of Web Users Based on Access Patterns," *Proc. of the ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 95-104.
- Huizingh, E.K.R.E.(2000), "The Content and Design of Web Site: An Empirical Study," *Information & Management*, vol. 37, pp. 123-134.

- Igor V.C., Scott G. and P. Smyth.(2000), "A General Probabilistic Framework for Clustering Individuals and Objects," *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 140-149.
- Jarvenpaa, S.L. and P.A. Todd(1997), "Consumer Reactions to Electronic Shopping on the World Wide Web," *International Journal of Electronic Commerce*, vol. 1, no. 2, pp.59-88.
- Kardaras, D. and B. Karakostas(1999), "The Use of Fuzzy Cognitive Maps to Simulate the Information Systems Strategic Planning Process," *Information and Software Technology*, vol. 41, pp. 197-210.
- Kim, H.S. and K.C. Lee(1998), "Fuzzy Implications of Fuzzy Cognitive Map with Emphasis on Fuzzy Causal Relationship and Fuzzy Partially Causal Relationship," *Fuzzy Sets and Systems*, vol. 97, pp. 303-313.
- Kosko, B.(1986), "Fuzzy Cognitive Maps," *International Journal of Man-Machine Studies*, vol. 24, pp. 65-75.
- .(1992), *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice-Hall.
- Kwahk, K.Y. and Kim, Y.G., "Supporting Business Process Redesign Using Cognitive Maps," *Decision Support Systems*, vol. 25, pp. 155-178. 1999.
- Larson, K. and M., Czerwinski(1998), "Web Page Design: Implications of Memory, Structure and Scent for Information Retrieval," *Conference on Human Factors and Computing Systems: CHI 98*, pp. 25-32.
- Lawrence, S. and C.L. Giles.(2000), "Efficient Identification of Web Communities Gary William Flake," *Proc. of the ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 150-160.
- Lee, S. and I., Han(2000), "Fuzzy Cognitive Map for the Design of EDI Controls," *Information & Management*, vol. 37, pp. 37-50.
- Liang, T.P. and H.J. Lai(2000), "Electronic Store Design and Consumer Choice: an Empirical Study," *Proceedings of the 33rd Hawaii International Conference on System Sciences: HICSS '2000*, pp. 1-10.
- Lightfoot, J.M.(1997), "A New Technique for Website Design Using an Interactive Visualization Hypertext Tool," *Proceedings of the 1998 conference on Computer personnel research SIGCPR '97*, pp. 14-19.
- Lohse, G.L. and P. Spiller(1998), "Quantifying the Effect of User Interface Design Features on Cyberstore Traffic and Sales," *Conference on Human Factors and Computing Systems: CHI '98*, pp. 18-23.
- Mannila, H. and C. Meek(2000), "Global Partial Orders from Sequential Data," *Proc. of the ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 161-168.
- Mannila, H., Toivonen, H. and I. Verkamo(1997), "Discovery of Frequent Episodes in Event Sequences," Department of Computer Science Series of Publications Report C-1997-15.
- McCallum, A., Nigam, K. and L.H. Ungar.(2000), "Efficient Clustering of High-Dimensional Data Sets with Application to Reference Matching," *Proc. of the ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 169-178.
- Mobasher, B., Cooley, R. and J. Srivastava.(1999), "Automatic Personalization Based on Web Usage Mining," *Proc. of the ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 123-132.

- Moe, W. and P.S. Fader(2000), "Capturing Evling Visit Behavior in Clickstream Data," *Working Paper*, The Wharton School, University of Pennsylvania.
- Montazemi, A.R. and D.W. Conrath,(1986), "The Use of Cognitive Mapping for Information Requirements Anslsysis," *MIS Quarterly*, vol. 10, no 1, pp. 45-56.
- Novak, T.P. and D.L. Hoffman(1997), "Measuring the Flow Experience Among Web Users," *Working paper*, Vanderbilt University.
- Park, J.S., Chen, M.S. and P.S. Yu(1995), "An Effective Hash-Based Algorithm for Mining Association Rules," *Proceedings of ACM SIGMOD*, pp. 175-186.
- Satur, R. and Z.Q. Liu(1999), "A Contextual Fuzzy Cognitive Map Framework for Geographic Information Systems," *IEEE Tractions on Fuzzy Systems*, vol. 7, no.5, pp. 481-494.
- Savasere A., Omiecinski E., and S. Navathe.(1995), "An Efficient Algorithm for Mining Ass ociation Rules in Large Databases," *Proceedings of the 21th International Conference on Very Large Databases*, pp. 432-444.
- Schneider, M., Schaidler, E., Kandel, A. and G., Chew (1998), "Automatic Construction of FCMs," *Fuzzy Sets and Systems*, vol. 93, pp. 161-172.
- Spiliopoulou, M., Pohle, C. and L.C. Faulstich(2000), "Improving the Effectiveness of a Web Site with Web Usage Mining," <http://www.wiwi.hu-berlin.de>.
- Srikant, R. and R. Agrawal(1995), "Mining Generalized Association Rules," *Proceedings of the 21th International Conference on Very Large Data-bases*, pp. 407-419.
- Srikant, R. and R. Agrawal(1996), "Mining Sequential Patterns: Generalizations and Performance Improvements," *Proceedings of the Fifth Int'l Conference on Extending Database Technology (EDBT)*, Avignon, France.
- Srivastava, J., Cooley, R., Deshpande, M. and P.N. Tan (1999), "Web Usage Mining: Discovery and Applications of Usge Paterns from Web Data," *Proc. of the ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1-10.
- Wells, J.D. and W.L., Fuerst(2000), "Domain-Oriented Interface Metaphors: Designing Web Interfaces for Effective Customer Inteeraction" *Proceedings of the 33rd Hawaii International Conference on System Sciences: HICSS '2000*, pp. 1-10.
- Zhang, P., Small, R.V., von Dran, G.M., and S. Barcellos (2000), "A Two Factor Theory for Website Design," *Proceedings of the 33rd Hawaii International Conference on System Sciences: HICSS '2000*, pp. 1-10.

Analysis of User's Purchasing Pattern on the Web Shopping Mall by Using Web Log Analysis and Fuzzy Cognitive Map Approach

Kun Chang Lee* · Soon Jae Kwon** · Jae Kyung Kim***

Abstract

A web shopping mall is a powerful vehicle for attracting a lot of customers to the Internet. To analyze customer's purchasing pattern, a wide variety of web mining tools were used in literature. However, those tools suffer from the lack of considering cause-effect relationship between several factors consisting of the web shopping mall. In this sense, we propose using fuzzy cognitive map(FCM) to analyze customer's purchasing pattern. To add more rigorous analysis procedures, we introduce a concept of core FCM and unified FCM. To induce a meaningful FCM, we applied web log analysis, sorting out a set of useful causal relationships among the components of web shopping mall. Simulation results with unified FCM yielded that the proposed approach using FCM is robust and powerful for analysing customer's hidden purchasing pattern on a specific web shopping mall.

Key words: Web log mining, Web log data, Causal relationships, Fuzzy cognitive map, Internet shopping mall, Consumer's behavior

* Professor of MIS School of Business Administration Sung Kyun Kwan University.

** Doctoral Candidate School of Business Administration Sung Kyun Kwan University.

*** Professor of MIS School of Business Administration Kyung Hee University.