한국지능정보시스템학회논문지 제13권 제1호 2007년 3월(pp.59~79)

P2P 환경에서 사회적 연결망을 활용한 지식관리시스템의 구축*

김윤상
㈜ 넷스루
(kimys@nethru.co.kr)

권순범
국민대학교 경영학부 e-비즈니스전공
(sbkwon@kookmin.ac.kr)

P2P(Peer to Peer)기술은 저비용과 편리성 등의 이유로 파일교환과 공유에 널리 활용되고 있다. 주로 검색 키워드와 파일의 이름을 매칭하여 P2P 네트워크에서 원하는 파일을 찾아주는 형태로 이용된다. P2P 환경의 가장 큰 특징은 네트워크 자체가 매우 동적으로 진화한다는 점이다. 즉, 노드의 추가와 삭제, 각 노드가 갖고 있는 파일의 변화, 네트워크의 변화에 따라서 키워드 검색 범위 등이 동적으로 변화한다.

본 연구는 P2P 환경에서 지식관리시스템(KMS: Knowledge Management System)의 구축을 제안한다. 제안하는 시스템은 지식이 네트워크의 여러 사람에게 광범위하게 흩어져 있고 각자가 보유한 지식도 변화하는 동적인 환경에서의 지식관리가 가능하며 P2P의 저비용과 편리성을 제공할 수 있다. 또한 기존의 P2P 기반 시스템은 단순 키워드 매칭만으로 검색하여 검색의 질이 낮았는데, 본 시스템에서는 사회적 연결망(Social Network) 기법을 활용하여 보유한 지식(관심도)이 가장 유사한 노드들을 논리적인 네트워크로 구성, 유지하여 지식의 검색과 관리의 효율을 높인다. 네트워크의 동적인 진화는 새로운 노드가 추가되거나 삭제될 때, 그리고 네트워크를 구성하는 각 노드가 새로운 지식(문서)을 추가한 경우(검색의 결과 새로운 문서를 다운로드한 경우 등) 논리적 네트워크의 변화로 진화하게 된다.

최근 인터넷에서는 수 많은 사이트(예: 블로그)에 지식이 흩어져 있고, 동적으로 변화하고 있다. 이러한 흩어진 지식을 어떻게 구조화하여 활용할 것인가가 중요한 이슈가 되고 있다. 기업 내부의 지식관리에 있어서도 여러 지역의 다양한 사람이 보유하고 있는 지식을 기업 지식포털로 통합하는 것이 큰 어려움이 중의 하나이다. 따라서 제안하는 지식관리시스템이 이러한 문제의 해결에 도움을 줄 수 있으리라 기대한다.

논문접수일 : 2006년 12월 게재확정일 : 2007년 03월 교신저자 : 김윤상

1. 서론

지식관리시스템(KMS: Knowledge Management System)은 조직원들 간의 지식을 수집, 저장, 배분하며, 새로운 지식의 창조를 지원하고, 개인의 잠재된 지식을 형상화하여 조직의 지식으로 승화 할 수 있도록 지원해 주는 정보시스템이다. 지식관리시스템은 조직원 각자의 관심과 필요에 맞는 전문지식을 찾는데 소요되는 시간과 노력을 절감 할 수 있게 하며, 문제 해결을 위해 전문적인 지식의

* 본 논문은 2005년도 국민대학교 교내연구비를 지원받아 수행한 연구임.

활용을 극대화하는 것이 목적이다.

한편, P2P(Peer to Peer)기술은 정보의 검색과 관리에 있어서 저비용 고효율성의 특징을 갖는 기술이며, 대표적인 P2P 프로그램으로 그누텔라(Gnutella), 냅스터(Napster), 소리바다, 당나귀(e-Donkey), 푸루나등이 있다. P2P 기술이 급성장한 배경은 개인 사용자들의 인터넷 대역폭의 확대와 초고속화이며, 사용자들 서로 간의 파일정보를 제공하고 받을 수 있는 파일공유(file sharing)기능 때문이다. 파일 정보의 공유 이외에도 P2P 기술을 기반으로 CPU 공유, 공동작업, 커뮤니케이션, 상거래 구현이 가능하다.

P2P 기술의 특징은 1) 네트워크를 구성하는 여러 노드에 정보가 흩어져 있는 분산 환경이라는 점, 2) 매번 검색에 있어서 노드가 보유한 정보, 접속여부, 접속 속도 등 노드의 변화된 상황을 반영한다는 점, 3) 네트워크의 모습이 계속 진화 한다는 점이다. P2P 기술은 네트워크의 역동성을 반영할 수 있는 유연성을 제공한다. 그러나, P2P 환경에서 지금까지의 검색 방법은 물리적 연결 네트워크를 따라서 이루어 졌고 관련성이 높은 정보를 가지고 있는 노드를 찾는 데에 있어서 정확성이 높지 않다.

정보와 지식이 네트워크를 구성하는 여러 사람에게 광범위하게 흩어져있고, 각자가 보유한 지식도 시시각각 변화하는 동적인 환경이 보편화 될 것으로 예상된다. Web 2.0의 사용자 생성 컨텐츠(UCC: User Created Contents), 블로그의 활성화, 태깅(Tagging) 기술 등이 이러한 변화를 반영하는 것들의 예이다. 기업에서도 정보와 지식이 지역적으로 광범위하게 흩어져 있고 각 지역에서의 정보와 지식이 변화하는 상황에서의 지식관리가 요구 되고있다.

본 논문은 P2P 기술의 장점을 살리면서 지식관리 목적에 활용할 수 있도록 사회적 연결망(SN: Social Network) 기법을 도입, 통합하였다. 따라서, 'P2P 환경에서 사회적 연결망을 활용한 지식관리 시스템(P2P-SN-KMS)'을 제안한다. P2P-SN-KMS 시스템은 물리적인 네트워크가 아닌 관심사가 유사한 노드를(비록 해당 노드가 물리적으로 멀리 떨어져 있더라도) 이웃노드로 설정하는 논리적 네트워크를 구성한다. 유사도는 각 노드가 보유한 자신의 관심사 키워드를 비교하여 측정한다. 이러한 논리적 네트워크는 사회적 연결망 기법에 의해 각 노드의 관심사 키워드의 변화, 검색 키워드와 검색 결과, 새로운 노드의 추가에 따라서 이웃노드의 평가와 갱신이 계속적으로 이루어지면서 네트워크가 진화한다.

논문의 구성은 다음과 같다. 제 2장에서는 P2P 환경에서 사회적 연결망을 활용한 지식관리 시스템에 대한 기존연구를 살펴보고, 제 3장에서는 P2P-SN-KMS 시스템의 구조와 기능을 설명한다. 제 4장에서는 P2P 환경에서 사회적 연결망을 활용한 지식관리 시스템의 예제를 살펴보고, 제 5장에서는 P2P-SN-KMS시스템의 성과를 검증하고, 마지막으로 제 6장에서는 결론 및 향후 연구 방향을 서술한다.

2. 문헌 연구

2.1 P2P(Peer to Peer)

P2P 서비스는 중앙 서버에 전적으로 의존하지 않고 PC간에 디지털 정보를 주고 받으며, 최소의 자원으로 많은 양의 데이터를 공유하는 시스템이다. P2P는 기술적인 측면에서 하이브리드 P2P와 순수 P2P로 구분되고, 하이브리드 P2P는 실제적

인 데이터 전송은 중앙서버가 존재하지 않지만, 정보관리 차원에서 중앙서버가 존재하는 구조이다. 하이브리드 P2P 시스템의 대표적 예는 소리바다, Napster이다. 순수 P2P 구조는 서버 없이 클라이언트만을 활용하여 서로간의 디지털 정보를 관리하고 전송하는 것이며, Gnutella가 대표적인 예이다.

<표 1> P2P 시스템 비교

특징 / 시스템	접속 방법	검색 방법	진화 방법
Gnutella	• 네트워크 연결을 위한 1개 이상의 이웃 IP • 특별한 이웃 형성 방법 없음	• TTL값을 적용한 Flooding 검색	• 없 음
소리바다	• 중앙 서버에 직접 연결	• 중앙 서버로부터 받은 사용자 IP 리스트로 직접 검색 요청	• 서버업데이트
Napster	• 중앙 서버에 직접 연결	• 로그인시 파일 인덱스 정보 전송 • 중앙서버에서 관리하는 파일 인덱스정보를 통하여 검색	• 서버업데이트

<표 1>은 대표적인 P2P 기술 기반의 파일공유 시스템인 Gnutella, 소리바다, Napster를 비교한 것이다. 소리바다와 Napster는 중앙집중식으로 중앙 서버가 존재하며, 검색 시 중앙서버에서 관리하는 정보를 이용하고 있다. Gnutella는 중앙 서버가 존재하지 않으며 최초 연결 시 반드시 네트워크에 연결된 1개의 다른 노드 IP주소를 알아야 하고, 검색 방법에서는 이웃노드를 통해 검색하는 Flooding방식을 사용한다. Gnutella는 특별한 진화 방법이 없으며, 중앙 서버가 통제하는 소리바다와 Napster는 중앙 서버의 업데이트로 네트워크가 진화한다[7].

2.2 SN(Social Network)

조직에서 개인의 사회적 관계는 조직원들간에 상호작용으로 어떤 문제를 해결하기 위해 협력 할 수 있는 개인들간의 관계이다. 개인이 문제를 해결하기 위해 가장 좋은 수단 중의 하나는 이미 같은 문제를 겪어본 사람을 찾아 조언을 듣는 것이다. 개인은 목적에 따라 많은 인간관계 네트워크를 형성하고 있으며, 네트워크를 활용하여 많은 문제를 해결하고 있다. 즉 개인이 풀어야 할 문제에 대하여 개인이 가지고 있는 인간관계 네트워크 망을 활용하는 것이다.

한편, 사회적 연결망 이론은 관련성이 높은 개별 행위자 사이에 형성되는 관계적 속성을 분석의 대상으로 삼고 있다. 개인의 상호작용 행위로 연결이 이루어지고 이러한 연결의 반복으로 네트워크가 역동적으로 만들어지고, 반대로 전체 네트워크의 형태가 개인의 상호작용 행위에 영향을 미친다는 이론이다. 이러한 이론을 구성원 간에 지식을 요청하고, 제공받고, 공유하는 지식 네트워크 형성과 관리에 적용 할 수 있다[4]. 이러한 상호 작용이 많은(관심사가 유사하여 지식의 교환이 많은) 사람에게 우선 지식을 요청하는 행위도 사회적 연결망 이론에 부합한다.

2.3 KMS(Knowledge Management System)

KMS는 개인과 조직간의 지식을 생성, 축적, 활용을 보다 원활하게 할 수 있도록 정보 기술을 통해 지원하는 것이다[6]. 개인 또는 조직원들이 사회 활동을 하면서 얻은 지식은 개인의 머릿속, 문

서, 파일, 이메일 등의 다양한 형태로 보유하고 있음에도 공간적, 시간적 제약으로 지식을 공유하기 어렵다. KMS는 지식을 디지털화하고, 네트워크를 이용하여 언제 어디서나 자유롭게 최적의 지식을 사용할 수 있도록 도와주는 시스템이다.

특히, 인터넷의 발달, 조직의 글로벌화, 조직구조의 분산화 등의 원인으로 지식의 생성과 축적이 여러 곳에 흩어져서 이루어지면서 과거에 비해서 상당히 역동적인 변화를 보인다. 따라서 KMS도 분산된 지식 중에서 관련성이 높은 지식을 선별적으로 찾아주고 역동적으로 변화하는 지식을 관리할 수 있어야 한다. 본 연구에서는 이러한 요구사항을 반영하기 위한 KMS를 제안하며, 지식관리의 대상을 문서와 문서의 키워드로 삼았다.

3. P2P-SN-KMS 시스템 구조와 기능

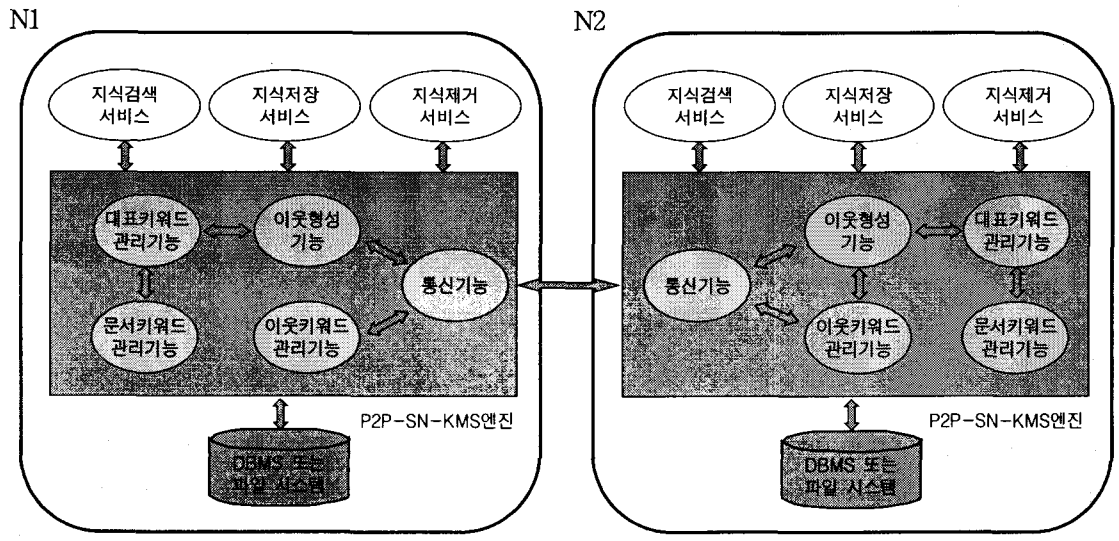
3.1 P2P-SN-KMS 아키텍처

P2P-SN-KMS의 아키텍처는 지식검색서비스, 지식저장서비스, 지식제거서비스, P2P-SN-KMS 엔진, DBMS(또는 파일시스템)으로 구성 되어 있다.

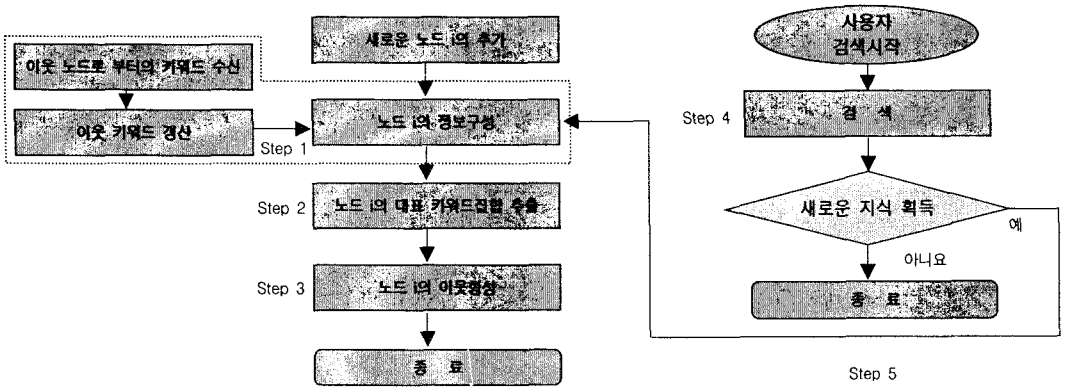
[그림 1]에서 지식저장, 지식제거, 지식검색 서비스는 사용자가 직접 요구하고 수행하는 서비스 기능이다. P2P-SN-KMS 엔진의 주요 역할은 노드가 보유한 지식의 관리와 논리적 지식네트워크의 관리를 위해서 문서키워드와 대표키워드 관리기능, 이웃형성과 이웃키워드 관리기능 그리고 다른 노드와의 통신기능을 가지고 있다. [그림 1]은 서로 같은 구조를 갖는 두 개의 노드 N1과 N2가 통신하는 예이다. [그림 1]에서 나타난 키워드 관련 용어는 3.2절의 Step 1에 설명되어 있다.

3.2 P2P-SN-KMS의 지식관리 기능

P2P-SN-KMS 시스템의 주요 지식관리 기능을 절차적으로 표현하면 [그림 2]와 같다. 네트워크의



[그림 1] P2P-SN-KMS 아키텍처



[그림 2] P2P-SN-KMS 기능(절차적 표현)

형성, 검색, 진화로 구성되어 있으며 각 기능은 다음과 같다.

네트워크 형성은 새로운 노드 i가 추가되면 지식관리를 위해서 노드 i와 관련된 정보를 획득하여 대표키워드를 추출하고, 추출된 대표키워드로 이웃을 형성하게 된다. 네트워크 형성 후 이웃노드를 활용하여 지식을 검색하고, 새로운 지식을 획득하면 이를 반영하여 노드 i의 정보구성을 업데이트 하고 대표키워드를 재 추출하여 이웃을 재설정하여 진화한다. <표 2>는 스텝 별로 기능 설명을 위한 수식에서 사용되는 변수와 인덱스를 정리 한 것이다.

<표 2> 수식 설명

기호	설명
N	네트워크
i (i = 1, 2, 3,…n)	노드
k (k = 1, 2, 3,…p)	키워드
d (d = 1, 2, 3,…m)	문서
RK_i	노드 i의 대표키워드 집합
Q	대표키워드 집합의 원소의 개수 (모든 노드가 동일)
NB_i	노드 i의 이웃노드 집합

Step 1: 노드의 정보 구성

각 노드는 네트워크 형성, 검색, 네트워크 진화를 위해 다음과 같은 정보를 갖는다.

① 문서키워드: 노드가 보유하고 있는 문서(지식)에서 추출 되어진 키워드이다(한 문서에 한 개 이상의 키워드가 있음)

② 대표키워드: 어느 노드의 전체 문서키워드 중에서 그 빈도가 높아서 노드의 주된 관심사라고 판단되는 키워드로, 모든 노드에 동일한 개수로 한정된다.

③ 이웃키워드: 어느 노드의 이웃으로 형성된(유사한 대표키워드를 갖는) 노드가 가지고 있는 키워드로, 이웃노드가 새로운 문서를 확보하면 이웃노드는 해당 문서의 키워드를 보내준다.

④ 각 노드 간의 연결을 유지 하기 위한 노드의 연결정보(IP번호와 Log-on/Log-off 상태)

자신과 다른 노드와의 유사도를 측정하기 위해서는 대표키워드를 추출해야 하며, 대표키워드 집

합의 추출과 이웃(논리적 네트워크) 형성 방법은 다음과 같다.

Step2 : 노드 i의 대표키워드집합 추출

Step2-1: 노드 i가 보유하고 있는 모든 문서 d(d=1,2,3,…m)에 등록된 키워드를 카운트 한다.

Step2-2: 키워드 k(k=1,2,3,…p)의 빈도를 계산한다.

Step2-3: 빈도 수가 높은 상위 Q개 키워드를 추출한다.

Step3: 노드 i의 이웃형성

Step3-1: P2P-SN-KMS 소프트웨어가 설치된 노드의 IP 정보를 이용하여, Flooding 방식으로 다른 노드 j의 대표키워드 요구

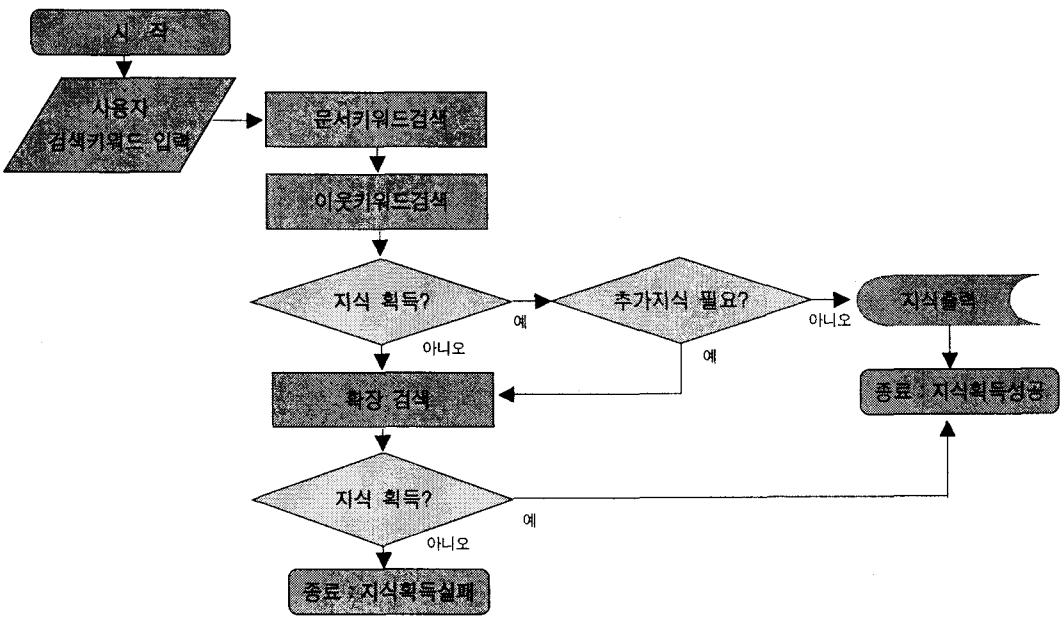
Step3-2: 노드 i와 노드 j의 대표키워드를 비교하여 공통의 키워드 수 카운트

Step3-3: 공통 키워드 수가 일정 수(0.2 * Q)보다 많은 노드를 노드 i의 이웃노드로 설정

$$NB_i = \{j \mid j \in N,\ RK_i \cap RK_j \neq \emptyset \text{ AND } n(RK_i \cap RK_j) \geq 0.2 * Q\}$$

Step4: 지식검색

지식검색은 3가지 검색 방법을 차례로 적용한다. ① 문서키워드 검색은 자신이 보유하고 있는 문서를 대상으로 한 검색이고, ② 이웃키워드 검색은 이웃노드가 지식을 가지고 있는가를 검색하는 것이고, ③ 확장 검색은 기존 P2P에서의 Flooding 방식과 동일한 것으로 이웃키워드 검색으로 지식을 찾지 못하였을 때, 이웃노드를 넘어선 확장된 범위에서 지식을 검색하는 것이다.



[그림 3] 지식검색

Step 5: 네트워크의 진화

P2P-SN-KMS는 다음의 두 가지 경우에 진화 과정을 거친다.

① 대표키워드의 변화: 각 노드의 대표키워드 집합의 변화로 이웃노드가 변화되는 경우로, 예를 들어 노드 i가 검색의 결과 새로운 지식(문서)을 획득하여 문서키워드 리스트가 변하면, 대표키워드 RK_i가 변화 할 수 있다.

Step5-1: 각 노드의 대표키워드 집합의 변화

Step5-2: 이웃노드 재설정(네트워크 진화)

② 새로운 노드의 추가: 새로운 노드가 네트워크에 추가되는 경우로, 새로 추가된 노드가 자신의 이웃을 형성하면서 네트워크가 진화하게 된다.

Step5-1: 새로운 노드가 네트워크에 진입

Step5-2: 새로 추가된 노드가 자신의 이웃을 형성(네트워크 진화)

3.3 기존 시스템과의 비교

<표 3>은 기존 시스템인 Gnutella(P2P), 사회적 연결망을 활용한 P2P 시스템(P2P-SN)과 본 논문이 제안 하는 시스템 P2P-SN-KMS을 비교한 것이다.

네트워크의 이웃형성방법, 형성된 네트워크에서 검색방법, 그리고 네트워크의 진화방법 세가지 관점에서 비교하였다. 첫째, 이웃 형성 방법에 있어서 Gnutella는 네트워크를 연결하기 위하여 1개 이상의 네트워크에 연결된 IP정보를 알아야 하며, 이것이 곧 이웃이 된다(물리적 연결). 반면에 P2P-SN와 P2P-SN-KMS에서는 자신과 다른 이웃들의 평가를 통하여 이웃을 형성한다(논리적 연결). P2P-SN은 자신이 보유한 목차 정보(자료)를 세분화하여 특정한 목차에 대해서 두 개 이상의 노드가 얼마나 관계가 있는가를 알아보고 가장 관계가 높은 노드를 이웃으로 설정한다[12]. 본 논문의 시스템인 P2P-SN-KMS는 자신이 보유한 모든 문서키워드를 기준으로 대표키워드를 추출하여 두 개 이상의 노드가 얼마나 유사한 대표키워드를 보유하고 있는가를 측정하여 이웃으로 설정한다.

<표 3> 기존 시스템 과의 비교

특징 / 시스템	이웃형성방법	검색방법	네트워크 진화방법
Gnutella(P2P)	• 특별한 이웃 형성 방법 없음 • 네트워크 연결을 위한 1개 이상의 이웃 IP	• TTL 값을 적용한 Flooding 검색	• 없음
P2P-SN (Upadrashta, 2005)	• 쿼리를 통한 관심그룹 형성	• 이웃 리스트를 통한 검색	• 새로운 이웃 발견
P2P-SN-KMS	• 문서키워드에서 대표키워드 추출 • 대표키워드를 통한 관심 그룹 형성	• 이웃키워드 검색 • Flooding 검색	• 새로운 대표키워드 추출

둘째, 검색 방법에 있어서 Gnutella는 TTL(Time To Live)를 적용한 Flooding 방법으로 검색을 시행한다[5]. P2P-SN는 앞에서 언급한 이웃을 형성한 리스트를 기반으로 이웃에게 검색을 요청한다. 본 논문의 시스템은 이웃노드가 보내주어 보유하고 있는 이웃키워드를 검색하여 이웃노드의 지식을 검색하고, 결과를 얻지 못하면 이웃의 경계를 넘어서 다른 노드를 Flooding 방식으로 검색한다.

셋째, P2P-SN의 네트워크 진화 과정은 기존의

네트워크 상태에서 새로운 노드가 추가되면 서로 간의 상호작용을 통해 자신과의 관계가 높은지를 판단하게 되고 기존의 이웃노드보다 관계가 더 높다면 이웃을 교체 하는 과정을 거친다[13]. 반면에 본 논문의 시스템은 자신이 보유한 문서키워드가 업데이트되면 대표키워드를 재 추출하게 되고, 이 대표키워드로 이웃을 재 설정하게 된다. 따라서 검색의 과정과 검색 종료 후 검색결과의 반영을 통해서 네트워크가 진화한다.

4. 예 제

<표 4>는 20개의 노드가 각각 3개의 대표키워드를 보유하고 있는(Q = 3) 예제이다(20개의 노드는 대표키워드 외에 다른 문서키워드는 보유하고 있지 않는다고 가정함). 본 예제를 대상으로 P2P-SN-KMS 시스템의 기능에 대해서 설명하고 검증한다.

4.1 이웃 형성

본 예제에서는 노드 1(앞으로 N1으로 표기)과 다른 노드간의 유사도를 측정하여 N1이 4개의 이웃노드를 형성 하는 과정을 표현한 것이다.

각 노드가 대표키워드를 N1에게 전달하면, N1은 각 노드와의 유사도를 계산한다. 유사도는 두 노드가 보유한 같은 키워드 개수로 RK_1 = {RFID, Smart Tag, MEMS}이고 N4, N7, N8, N11의 대표키워드와 두 개가 일치하고, N2, N12, N16, N17, N18, N20의 한 개의 대표키워드가 일치 한다. 따라서, 유사도가 높은 노드는 N4, N7, N8, N 11이며, 네 개의 노드가 이웃노드로 형성 된다. N1은 4개의 이웃노드를 형성 하였으므로, N1의 이웃키워드는 각 이웃노드의 대표키워드인 N4(Smart Tag, RFID, FTP), N7(인증시스템, Smart Tag, RFID), N8(ISMS, MEMS, RFID), N11(RFID, Smart Tag, 다운사이징)이 이웃키워드가 된다.

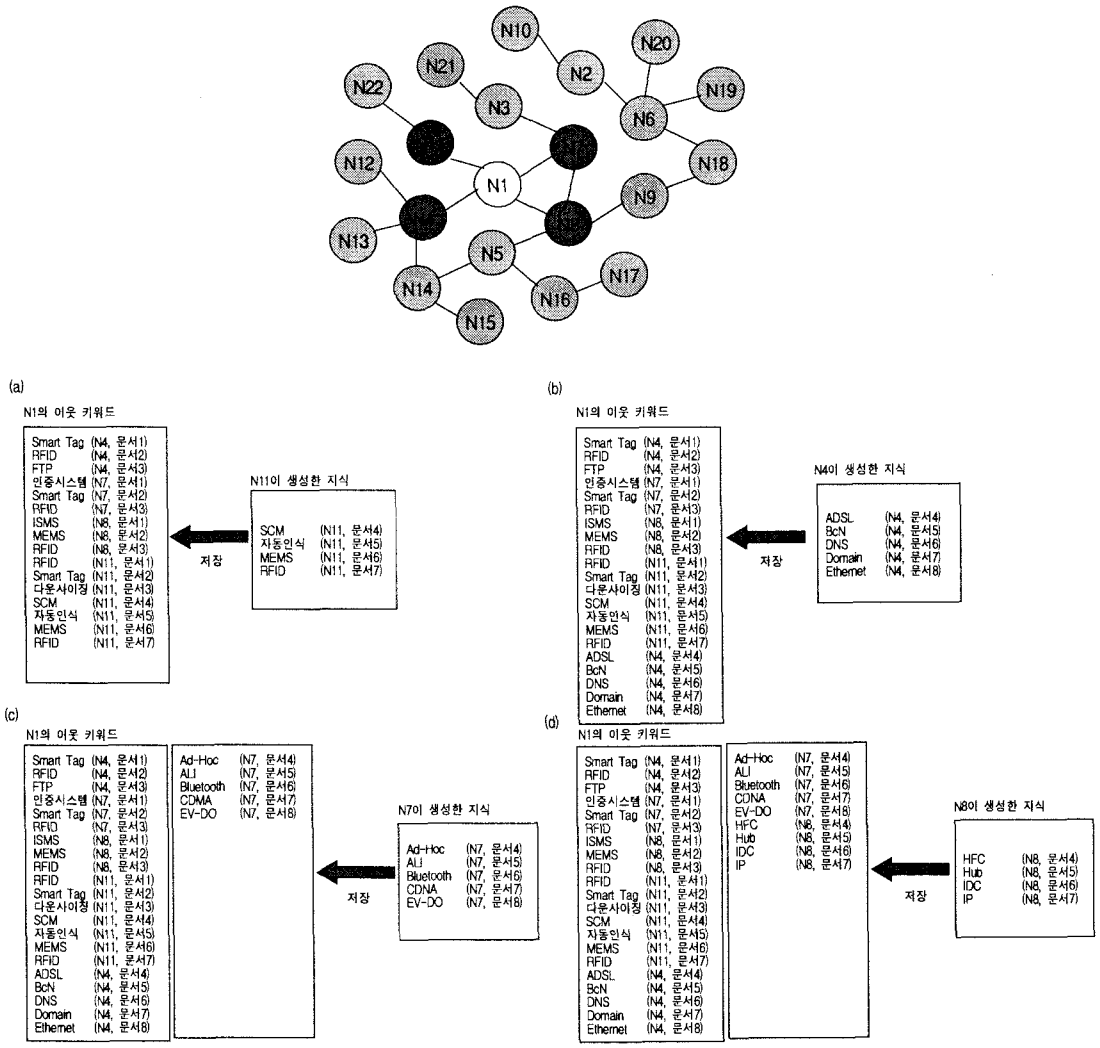
<표 4> 대표키워드

명칭	대표 키워드	명칭	대표 키워드
N1	RFID	N11	RFID
N1	Smart Tag	N11	Smart Tag
N1	MEMS	N11	다운 사이징
N2	E2E	N12	E2E
N2	RFID	N12	RFID
N2	Embedded System	N12	FTP
N3	자동인식	N13	GPS
N3	다운 사이징	N13	다운 사이징
N3	전자봉투	N13	IPv6
N4	Smart Tag	N14	CDMA
N4	RFID	N14	ISMS
N4	FTP	N14	FTP
N5	HFC	N15	HFC
N5	Napster	N15	IP
N5	GPS	N15	인증시스템
N6	IP	N16	GPS
N6	ISDN	N16	전자봉투
N6	인증시스템	N16	RFID
N7	인증시스템	N17	ISMS
N7	Smart Tag	N17	Smart Tag
N7	RFID	N17	다운사이징
N8	ISMS	N18	ISMS
N8	MEMS	N18	MEMS
N8	RFID	N18	HFC
N9	ISMS	N19	E2E
N9	IP	N19	ISDN
N9	IPV6	N19	Napster
N10	Cookies	N20	자동인식
N10	CDMA	N20	FTP
N10	GPS	N20	MEMS

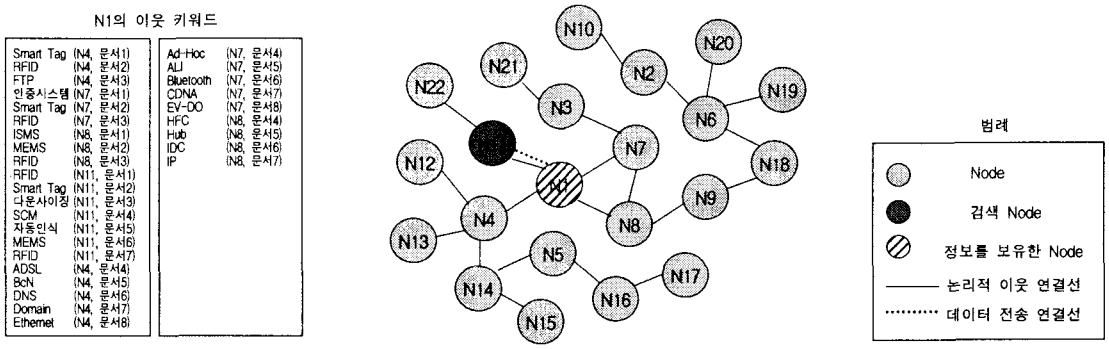
4.2 지식검색

N1의 대표키워드 비교로 이웃으로 형성된 N4, N7, N8, N11이 새로운 지식을 추가, 삭제를 하게 되면 즉, 이웃노드의 문서키워드의 변화가 생기면 N1의 이웃키워드가 갱신된다. 이웃키워드가 갱신 또는 추가 될 때 이웃키워드가 어디서 생성 되었

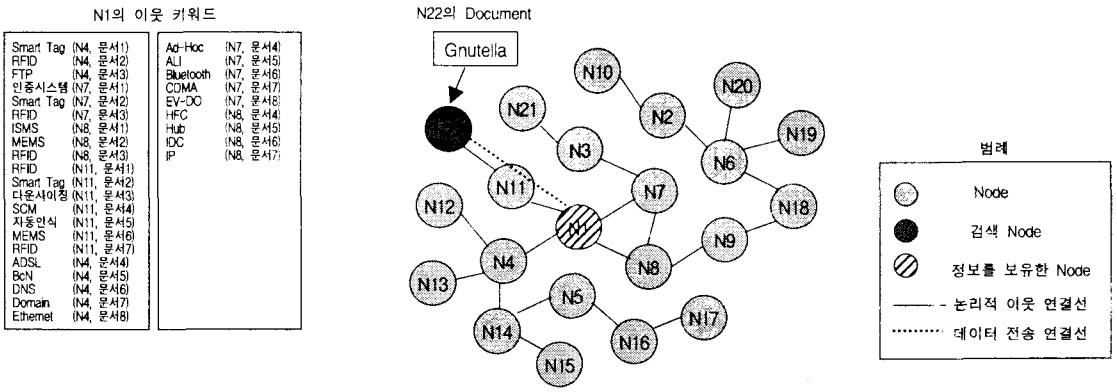
는지 알 수 있는 소스 주소와 어느 문서에서 추출된 이웃키워드인지 구분하는 문서 정보를 담고 있다. (a)에서 N11은 SCM, 자동인식, MEMS, RFID 지식을 새롭게 획득하여 문서키워드를 추가하면, N1은 N11로부터 새로운 이웃키워드, 이웃키워드 소스 주소, 이웃키워드 문서정보를 추가한다. N1에 있는 기존의 이웃키워드는 N1의 이웃 형성 시에 이웃노드 N4, N7, N8, N11로부터 전달 받은 각각의 대표키워드다. 마찬가지로(b)는 기존에 지식이 담긴 이웃키워드에 추가적으로 N4가 생성한 지식 ADSL, BcN, DNS, Domain, Ethernet을 생성으로 문서키워드가 추가되면 N1의 새로운 이웃키워드, 이웃키워드 소스 주소, 이웃키워드 문서정보가 추가된다. (c)와 (d)도 같은 방법으로 N7번, N8번이 지식을 생성하면 N1의 이웃키워드, 이웃키워드 소스 주소, 이웃키워드 문서정보가 추가 된다.



[그림 4] 이웃키워드 생성



[그림 5] 이웃키워드 정보 검색



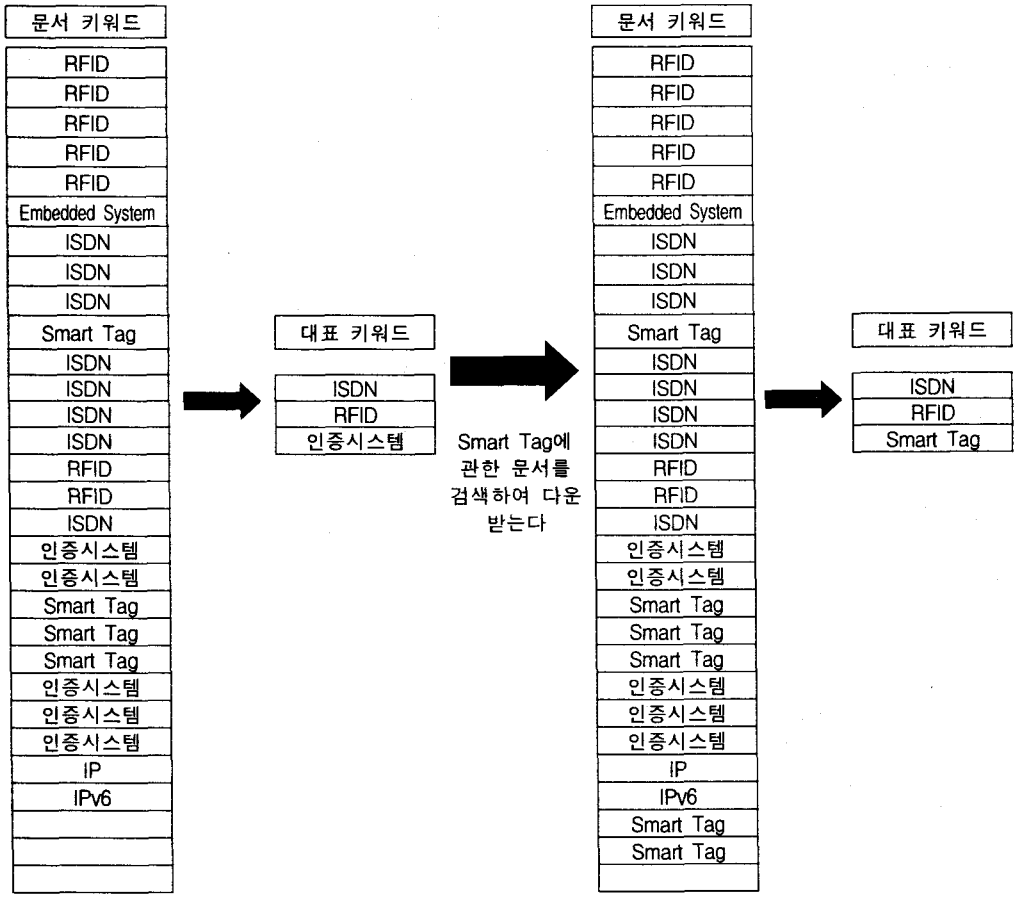
[그림 6] 확장검색

이후 N1에서 SCM을 검색하게 되면 처음 자신의 문서키워드에 정보가 있는지를 확인하고 없으면, 이웃키워드를 검색한다. [그림 5]는 자신의 이웃키워드에 SCM이 있으므로, [그림 4]에서 N11이 생성한 SCM의 지식을 직접 연결하여 문서를 다운로드 한다.

[그림 6]은 사용자가 Gnutella라는 검색을 입력한 경우다. N1은 자신의 이웃키워드에 정보가 있는지를 확인하고 정보가 없으면, [그림 6]처럼 N1은 검색을 위한 쿼리문을 작성하고, 이 쿼리문은 이웃노드인 N4, N7, N8, N11에게 전달된다. 쿼리를 전달 받은 노드는 자신의 이웃키워드를 검색하고 다시 이웃에게 쿼리를 전달한다. 이것은 기존의 검색방법인 Flooding 방식과 유사하나 네트워크의 물리적 연결이 아닌 논리적연결(보유한 지식의 유사성이 높은 이웃노드 연결)을 따라서 검색을 수행한다. 이처럼 N1은 N11을 통해 N22에게 Gnutella라는 지식이 있다는 것을 알게 되고, 이 자료를 받기 위해 직접 N22와 연결하여 문서를 다운 받게 된다.

4.3 네트워크의 진화

모든 노드는 최초 이웃을 설정한 이후 계속적으



[그림 7] 대표키워드 변화

로 진화하게 된다. 본 시스템은 대표키워드가 교체되면 네트워크가 진화한다. [그림 7]은 문서키워드가 업데이트 되는 과정에서 대표키워드 교체가 일어난 경우의 예이다.

최초 대표키워드는 보유하고 있는 문서키워드를 기준으로 대표키워드를 선정하게 된다. 예제에서는 이해를 돕기 위해 대표키워드가 3개라고 가정한다. [그림 7]에서 보면 최초 키워드는 가장 빈도수가 많은 ISDN(8개), RFID(7개), 인증시스템(5개)로 선정 되었다. 이 후 사용자는 계속적인 검색을 통하여 Smart Tag에 관한 문서를 2개를 다른 노드에서 다운 받게 되면, ISDN(8개), RFID(7개), Smart Tag(6개), 인증시스템(5개) 가 되어 대표키워드가 ISDN, RFID, Smart Tag으로 교체가 이루어진다. 교체가 이루어 지게 되면 이웃노드를 포함하여 자신과 유사한 노드를 다시 찾게 되며, 새로운 이웃을 형성하여 네트워크의 논리적 연결이 새롭게 진화한다. 만약 문서키워드 정보는 갱신 되었으나 대표키워드 교체가 이루어지지 않으면 진화를 하지 않는다. 따라서 한 노드에 지식이 적을 경우 대표키워드 교체가 빈번히 일어나 계속적인 진화과정을 거치게 되고, 거듭된 진화과정에서 문서

키워드 정보가 많이 쌓이게 되면 대표키워드 교체가 거의 일어나지 않게 된다.

P2P-SN-KMS 시스템은 JAVA를 이용하여 프로토타입 시스템을 구축하였다. 프로토콜로는 HTTP로 웹 기반으로 구축 하였다. P2P-SN-KMS 시스템의 효율성 검증을 위해 시뮬레이션을 수행하였다.

5. 검증

5.1 실험 방법

본 논문의 검증을 위해 Gnutella, P2P-SN, P2P-SN-KMS 세 가지 시스템에 대해 시뮬레이션을 수행 하였다. 시뮬레이션에서 3개 시스템의 검색 성능을 비교하였는데, 검색의 성능은 찾은 키워드 수, 방문한 노드수, 찾는데 걸린 시간, 찾은 키워드 수/방문한 노드 수(P)로 측정하였다. '방문한 노드 수'는 3개 시스템의 검색 과정에서 이웃으로 찾은 노드 수로, 중복되게 이웃으로 설정된 노드는 하나의 노드로 카운트 하였다(실제로 검색 과정에서 이미 방문한 노드를 중복 방문하지 않는다). '찾은 키워드 수'는 검색 하고자 하는 키워드(예를 들어 SCM)를 방문한 노드에서 얼마나 많이 찾았는가를 카운트 한 것으로 하나의 노드에 같은 키워드를 여러 개 가질 수 있다. 검색 성능을 나타내는 찾은 키워드수, 방문한 노드 수, 탐색 시간, P값(찾은 키워드 수/찾은 노드 수)은 모두 10번의 검색을 수행한 평균 값으로 구했으며, 10번의 검색은 네트워크를 구성하는 노드 중에서 10개의 노드를 임의로 선정 하였다. 전체 네트워크의 노드 수는 200개이고, 시뮬레이션은 다음의 5가지 변수에 대해서 수행하였다.

① TTL(Time To Live) : 검색을 수행한 노드에서부터 논리적으로(Gnutella의 경우는 물리적) 얼마나 멀리까지 검색을 수행하는가로 TTL 값이 클수록 확장검색의 범위가 더 넓어진다.

② 대표키워드 수(Q): P2P-SN-KMS 시스템에만 해당되는 변수로 각 노드가 보유하는 대표키워드의 개수

③ 노드가 보유하는 최대 문서키워드 수: 각 노드가 보유한 문서키워드 수가 많을수록 모든 노드가 지식을 많이 확보한 상태를 의미한다.

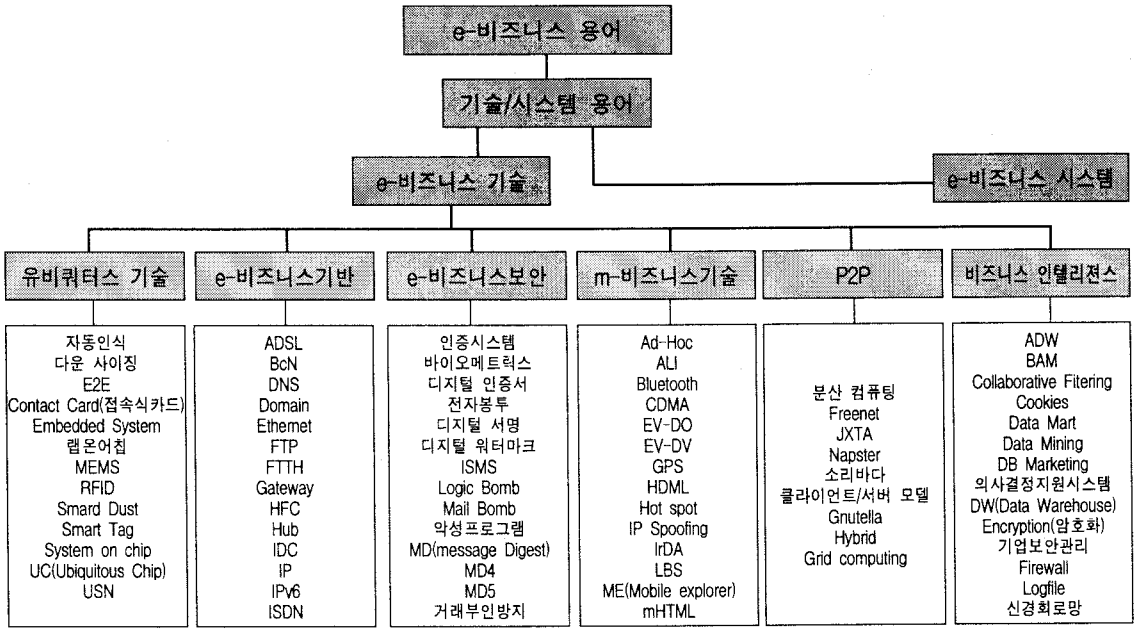
④ 노드가 이웃으로 형성 할 수 있는 이웃노드의 수: 각 노드는 이웃노드 수가 많을수록 확장검색을 하지 않고도 많은 검색을 할 수 있다.

⑤ 진화 메커니즘의 여부: P2P-SN-KMS 시스템은 검색의 결과로 인해 대표키워드의 변화 → 논리적 이웃노드의 변화 → 네트워크의 변화가 유발되는 진화 과정을 거친다. 진화 이전과 진화된 이후에 검색의 효율을 비교한다.

시뮬레이션 수행에서 시스템 성과에 영향을 미치는 주요 환경변수는 다음의 4가지로, 4종류의 시뮬레이션에서 이들 환경 변수를 고정시켜 성능 실험을 하였다.

<표 5> 환경 변수

변 수	설 명
대표키워드 수	P2P-SN-KMS의 대표키워드 수
보유키워드 수	각 노드가 최대로 가질 수 있는 문서키워드 수(최소 값은 대표키워드 수)
이웃노드 수	Gnutella, P2P-SN, P2P-SN-KMS가 이웃으로 형성 할 수 있는 최대의 이웃노드 수
탐색 시간	문서키워드를 찾는데 걸린 시간



[그림 8] 키워드 종류

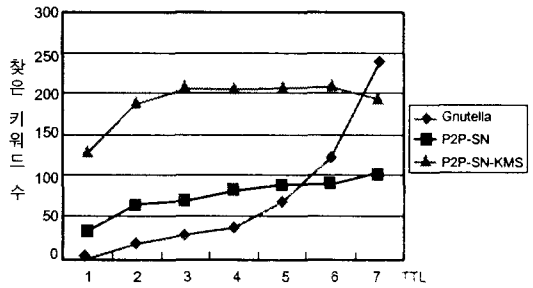
대표키워드 수는 P2P-SN-KMS에서만 사용되는 키워드 이고, 보유키워드 수는 각 노드가 최대 보유 할 수 있는 문서키워드이다. 예를 들어 대표키워드 수가 20이고 보유키워드 수가 200이라고 가정하면, 시뮬레이션을 위해 무작위로 각 노드는 20~200개의 키워드를 추출하고, 이때 추출하는 키워드의 70%는 자신의 관심목차에서 추출하고 나머지 30%는 다른 목차에서 추출한다. [그림 8]은 시뮬레이션에서 사용된 키워드로, 한국전자거래진흥원에서 발간한 "Dictionary of e-Business"에서 추출하였고, 6개 그룹 중에서 한 그룹을 선택하여 자신의 관심 목차로 삼는다. 이웃노드 수는 각 시스템이 최대로 맺을 수 있는 이웃노드의 수이다. 마지막으로 모든 실험에서 각 노드가 검색키워드를 선정 할 때는 자신의 관심목차에서 무작위로 1개를 추출하여 검색하며, 매 실험 마다 자신의 관심사 목차는 일정하다.

5.2 TTL의 변화

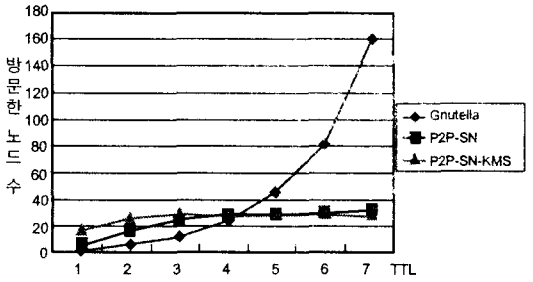
TTL 값의 변화로 3개의 시스템이 얼마나 많은 키워드를 찾았는지를 분석함으로써 시스템의 성능을 객관적으로 분석하였다. 시뮬레이션에서 고정시킨 변수 값과 환경설정은 다음과 같다.

① P2P-SN-KMS의 대표키워드 수는 20개.
② 노드가 가질 수 있는 키워드는 20~200개까지 무작위로 추출 한다.
③ 각 노드가 키워드를 추출 할 때 [그림 8]에서 무작위로 6개의 그룹 중 1개의 그룹을 선택하여 관심목차로 삼고, 70%는 관심목차에서 30%는 그 외 목차에서 무작위로 키워드를 추출 한다.
④ 각 노드가 이웃으로 형성 할 수 있는 최대의 이웃은 5개이다.

변수로는 TTL 값을 1부터 6까지 변화시켜 3개의 시스템을 비교하여 결과를 산출 하였고, 아래 [그림 9], [그림 10]은 측정한 결과 값의 그래프이다.

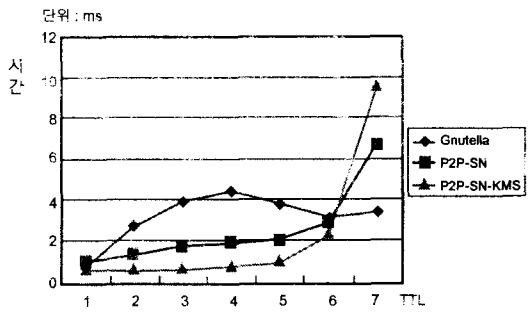


[그림 9] TTL 값에 따른 검색성능(찾은 키워드 수)



[그림 10] TTL 값에 따른 검색성능(방문한 노드 수)

[그림 9]와 [그림 10]에서 보면 TTL 값이 4부터 P2P-SN, P2P-SN-KMS가 찾은 키워드 수가 크게 증가 하지 않는 것으로 나왔다. 그 이유는 [그림 10]을 보면 TTL 값이 4일 때 P2P-SN, P2P-SN-KMS가 방문한 노드 수가 29정도로 일정한 수준에 이르렀기 때문이다. 반면에 Gnutella는 급격하게 상승 하는 곡선을 보이고 있는데, Gnutella는 이웃을 선정 할 때 무작위로 네트워크에 연결된 1~5개의 노드를 이웃으로 형성하기 때문에 TTL 값이 증가하면 찾는 이웃노드 수가 급격히 상승하게 되고, 이로 인하여 찾는 키워드 수도 상승하게 되어 TTL 값이 7에서는 P2P-SN-KMS보다 더 많은 키워드를 찾게 되었다. 하지만 TTL 값이 7에서 보면 P2P-SN-KMS, P2P-SN, Gnutella는 각각 방문한 노드 수가 약 28, 32, 160이고, 찾은 키워드 수는 각각 약 195, 100, 245이다. 검색의 TTL 값이 7일 때, 효율성을 보면 P2P-SN-KMS는 1개의 노드에서 약 6.9개 키워드를 찾았고, P2P-SN은 약 3.1개, Gnutella는 약 1.5개로 P2P-SN-KMS가 가장 효율이 좋았다. 더불어 [그림 10]에서 보면 TTL 값이 4일 때부터 P2P-SN-KMS, P2P-SN은 이웃 노드 수가 약 30개에서 크게 증가하지 않았다. 이것은 TTL 값이 4일 때 관심 그룹이 형성된 대부분의 노드를 찾았다는 것으로 확인 할 수 있다.



[그림 11] 1개의 키워드를 찾는데 걸린 시간

[그림 11]은 각 시스템이 TTL 값 변화에 따라 1개의 키워드를 찾는데 걸린 시간을 비교한 그래프이고, Y축의 시간의 단위는 ms(1/1000초)이다. Gnutella는 TTL값에 관계없이 그래프가 불규칙한 것을 볼 수 있는데 그 이유는 Gnutella가 무작위로 이웃을 형성하기 때문에 각 노드마다 또는 TTL값의 변화, 네트워크가 어떻게 연결되어 있는지에 따라서 다소 차이가 생긴다. P2P-SN과 P2P-SN-KMS는 점증적으로 1개의 키워드를 찾는데 시간이 증가하는 것을 볼 수 있는데 특히

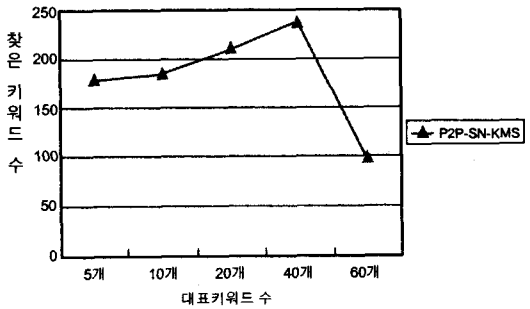
TTL 값이 5보다 커지면 급격히 시간이 증가하는 것을 볼 수 있다. 이것은 두 개의 시스템이 서로 유사한 이웃을 형성하기 때문에 2개의 서로 다른 노드가 같은 노드를 이웃으로 형성하고 있을 때 같은 노드에게 중복된 전달(쿼리)을 제거 하는데 걸리는 시간이 증가 하기 때문이다. 즉 TTL 값이 7일 때를 보면 실제적으로 키워드를 찾는 시간 보다 중복된 전달(쿼리)를 제거 하는데 더 많은 시간이 소비 된다는 것이다. P2P-SN-KMS는 방문한 이웃노드 수와 찾은 키워드 수가 TTL 값이 4일 때부터는 거의 증가 하지 않았고, 탐색 시간도 짧은 것을 확인 하였다.

5.3 P2P-SN-KMS의 대표키워드 수의 변화

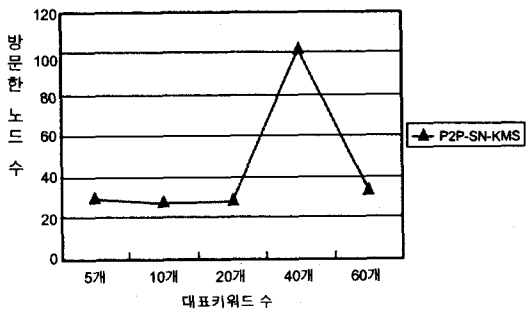
P2P-SN-KMS의 대표키워드 수를 변화 시키면서 분석한 데이터는 P2P-SN-KMS가 키워드 종류 수에 비례하여 몇 개의 대표키워드를 선출 하였을 때 가장 성능이 뛰어 났는지를 알아 보는 것이다. 이 분석 데이터는 현실에서 각 노드가 검색을 통하여 얻어진 문서키워드 종류에 비례하여 자신의 대표키워드를 몇 개를 추출해야 가장 좋은 성능을 낼 수 있는지를 판단하는 기준이 된다. 다음은 성능을 분석하는 데에 있어 환경 변수의 상수 값이다.

① TTL 값은 5.2절에서 가장 성능이 좋았던 값인 4로 선정하였다.
② 노드가 가질 수 있는 키워드는 20~200개까지 무작위로 추출 한다.
③ 각 노드가 키워드를 추출 할 때 [그림 8]에서 무작위로 6개의 그룹 중 1개의 그룹을 선택하여 관심목차로 삼고, 70%는 관심 목차에서 30%는 그 외 목차에서 무작위로 키워드를 추출 한다.
④ 각 노드가 이웃으로 형성 할 수 있는 최대의 이웃은 5개이다.

변수로는 대표키워드 수를 5~60개까지 선출하여 비교하였고. [그림 12], [그림 13]은 분석한 데이터에 그래프이다.



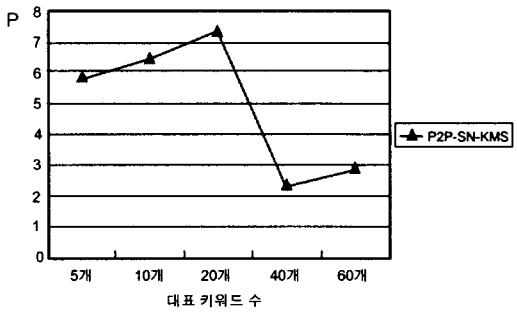
[그림 12] 대표키워드 수 값에 따른 찾은 키워드 수



[그림 13] 대표키워드 수 값에 따른 방문한 노드 수

[그림 12]와 [그림 13]을 보면, P2P-SN-KMS의 대표키워드 수가 5~20 범위에서 방문한 노드 수는 거의 일정하였고, 찾은 키워드 수는 대표키워드가 40개 일 때 가장 많았고 방문한 노드 수도 100개로 급격히 증가 하였다. TTL 값이 일정한 가운데 방문한 노드 수가 적다면 서로간의 관심 그룹이 잘 형성 되었다고 할 수 있으므로, 대표키워드 수가 20개 일 때 1개의 노드에서 약 7.1개의 키워

드를 찾았고 대표키워드 수가 40개 일 때 1개의 노드에서 약 2.4개의 키워드를 찾았다. 위 실험에서는 P2P-SN-KMS의 대표키워드 수가 20개(약 25%)일 때 가장 좋은 성능을 보였다. [그림 14]는 대표키워드 5~20개 까지는 P값이 상승하며 대표키워드 수가 20개 일 때 가장 효율이 좋은 것으로 확인 하였고 대표키워드가 20개를 넘으면 P값이 줄어들어 성능이 떨어지는 것을 알 수 있었다.



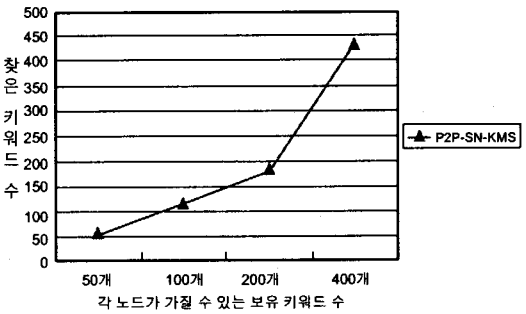
[그림 14] 대표키워드 수 값에 따른 P값

5.4 보유 키워드 수의 변화

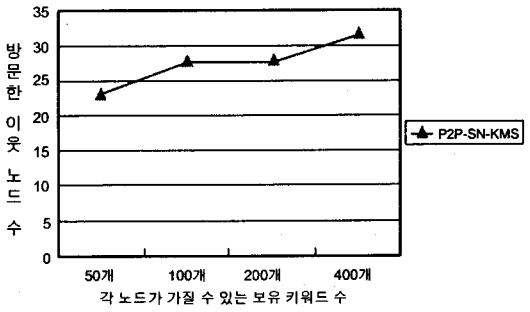
각 노드가 보유 키워드 수를 변화 시키면서 측정한 데이터는 시간의 관점에서 보았을 때 초기 네트워크 상태와 오래 지속된 네트워크 상태를 비교 할 수 있는 분석 데이터이다. 즉, 각 노드가 보유한 키워드 수가 적은 것은 초기 네트워크 상태에서 각 노드가 보유한 문서(지식)가 많지 않은 상황을 가정 할 수 있고, 반대로 보유한 키워드 수가 많은 것은 네트워크가 오래 지속되어 각 노드가 많은 문서를 보유한 상태로 가정 할 수 있다.

① TTL 값은 5.2절에서 가장 성능이 좋았던 값인 4로 선정하였다.
② P2P-SN-KMS의 대표키워드 수는 5.3절에서 효율이 가장 좋았던 20개로 정하였다.
③ 각 노드가 키워드를 추출 할 때 [그림 8]에서 무작위로 6개의 그룹 중 1개의 그룹을 선택하여 관심목차로 삼고, 70%는 관심 목차에서 30%는 그 외 목차에서 무작위로 키워드를 추출 한다.
④ 각 노드가 이웃으로 형성 할 수 있는 최대의 이웃은 5개이다.

변수로는 각 노드가 보유 할 수 있는 키워드 수 20~50, 20~100, 20~200, 20~400개로 각각 다르게 실험 하였다.



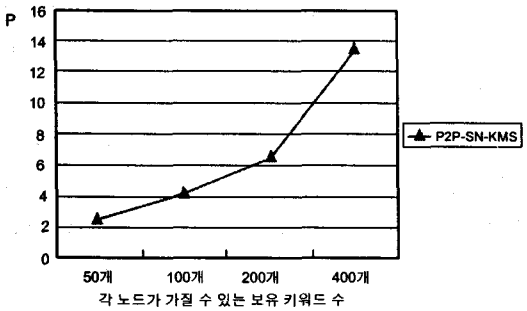
[그림 15] 보유 키워드 수 변화에 대한 찾은 키워드 수



[그림 16] 보유 키워드 수 변화에 대한 방문한 노드 비교

[그림 15]에서 P2P-SN-KMS는 보유 키워드 수에 비례하여 더 많은 키워드를 찾는 것으로 확인 할 수 있었다. [그림 16]는 P2P-SN-KMS의 방문

한 노드 수가 크게 변하지 않았다. 방문한 노드 수가 크게 변하지 않는 다는 것은 네트워크가 성장기에 이르렀을 때에도 자신들의 관심도가 높은 이웃끼리 잘 형성 된다는 것이다.



[그림 17] 보유 키워드 수 변화로 P값 비교

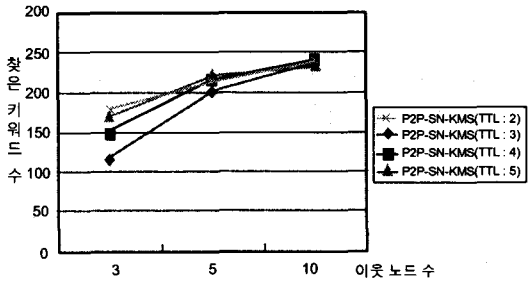
[그림 17]은 보유 키워드 수 변화에 따른 P값을 비교 한 것이다. P2P-SN-KMS는 보유 키워드 수가 증가 함에 따라 P값이 상승하는 것을 확인할 수 있다. 위 실험에서는 P2P-SN-KMS가 네트워크 생성의 초기 단계부터 성숙기까지 점증적으로 시간이 흘러도 성능을 유지 한다고 볼 수 있다.

5.5 이웃노드 수 변화

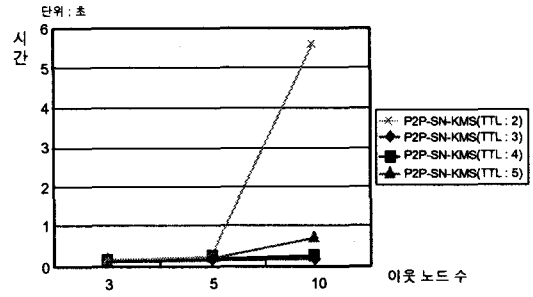
이웃노드 수를 변화시킴으로써, 얼마나 좋은 효율을 나타냈는지 알아보는 실험이다. 이 실험은 실제 네트워크 상에서 노드가 몇 개의 이웃노드를 형성하였을 때 가장 좋은 효율을 보였는지에 대한 실험이다. 실험을 위하여 상수 값은 다음과 같이 정했다.

① 노드의 수는 200개.
② TTL 값은 2, 3, 4, 5를 비교하였다.
③ P2P-SN-KMS의 대표키워드 수는 5.3에서 효율이 가장 좋았던 20개로 정하였다.
④ 노드가 가질 수 있는 키워드는 20~200개까지 무작위로 추출 한다.
⑤ 각 노드가 키워드를 추출 할 때 [그림 8]에서 무작위로 6개의 그룹중 1개의 그룹을 선택하여 관심 목차로 삼고, 70%는 관심 목차에서 30%는 그외 목차에서 무작위로 키워드를 추출 한다.

변수로는 각 노드가 이웃으로 형성 할 수 있는 이웃노드 수를 3~10개로 변화를 주었다.



[그림 18] 이웃노드 수 값에 따른 찾은 키워드 수 비교



[그림 19] 이웃노드 수 값에 따른 검색시간 비교

[그림 18]과 [그림 19]는 이웃노드 수 변화에 따라 찾은 키워드 수와 검색시간을 TTL 값 2, 3, 4, 5를 비교 한 것이다. 대체적으로 이웃노드 수가 증가하면, 찾는 키워드 수가 점증적으로 증가 하였으나, [그림 19]에서 보는 바와 같이 이웃노드 수가

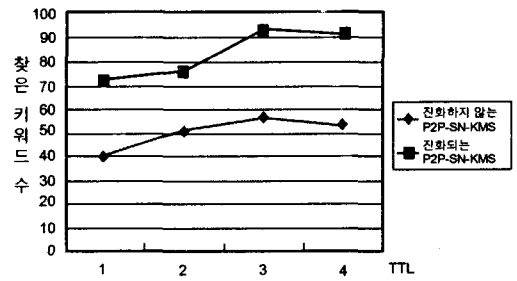
10개 일 경우 검색 시간은 이웃노드 수가 5일 때보다 최소 3배에서 최대 30~40배 이상 많은 시간이 소요되었다. 이것은 P2P-SN-KMS가 너무 많은 이웃을 허용하게 될 경우 찾는 지식의 양에 비해 찾는 시간이 더 많이 걸리고, 네트워크 상에서 많은 양의 트래픽을 유발하게 된다. 위 실험에서 TTL 값 2, 3, 4, 5일 때 거의 유사한 결과를 얻었으나, 그 중 가장 효율이 좋았다고 판단 할 수 있는 것은 TTL 값이 4이고, 이웃노드 수가 5일 때 P2P-SN-KMS가 가장 효율이 좋았다.

5.6 TTL 값을 기준으로 P2P-SN-KMS의 진화 성능 비교

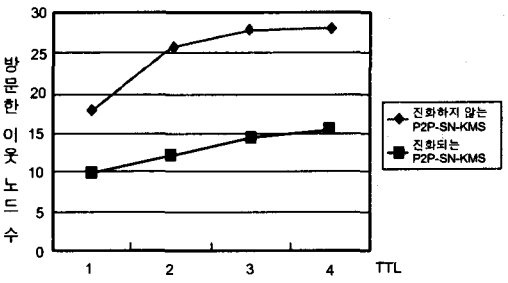
P2P-SN-KMS는 항상 동적으로 자신의 문서키워드를 기반으로 대표키워드를 추출하면서 새로운 지식의 습득할 때 상황에 맞추어 이웃노드를 변경하게 된다. 성능 비교 실험은 각 노드가 계속적으로 지식을 습득하여 문서키워드가 추가 될 때, 이를 반영하여 이웃노드를 바꾸는 P2P-SN-KMS (진화하는 경우)와 이를 반영하지 않는 P2P-SN-KMS (진화하지 않는 경우)를 비교 실험해 보았다.

① 최초 키워드 선출 150개(네트워크 진화 전).
② 최초 키워드 150개 선출 후, 100개 추가 선출 (네트워크 진화 후).
③ TTL 값은 5.2절에서 가장 성능이 좋았던 값인 4로 선정하였다.
④ P2P-SN-KMS의 대표키워드 수는 5.3절에서 효율이 가장 좋았던 20개로 정하였다.
⑤ 각 노드가 키워드를 추출 할 때 [그림 8]에서 무작위로 6개의 그룹 중 1개의 그룹을 선택하여 관심목차로 삼고, 70%는 관심 목차에서 30%는 그 외 목차에서 무작위로 키워드를 추출 한다.
⑥ 각 노드가 이웃으로 형성 할 수 있는 최대의 이웃은 5개이다.

최초 P2P-SN-KMS가 150개의 키워드를 선출하여 이웃으로 형성된 이웃노드 정보와 추가로 100개의 키워드를 선출한 후 이웃으로 형성된 정보를 다르게 하였다. 처음 150개를 선정할 때 70%로 선출된 1차 관심 목차를 제외하고 5개의 목차에서 1개의 2차 관심 목차를 임의적으로 선택하여 70%는 2차 관심목차에서 나머지 30%는 5개의 다른 목차에서 키워드를 무작위로 추출 하였다. 성능 비교를 위해 각 노드에 검색 키워드는 2차 관심 목차에서 무작위로 추출 하였다.



[그림 20] 진화기능 여부에 대한 찾은 키워드 수



[그림 21] 진화기능 여부에 대한 이웃노드 수

[그림 20]과 [그림 21]을 보면 양 쪽 모두 TTL 값에 따라 찾은 키워드가 증가 하였다. [그림 20]

에서 진화하는 P2P-SN-KMS가 더 많은 키워드를 찾았으며, 방문한 이웃노드 수는 적은 것을 알 수 있다. 진화기능이 없는 P2P-SN-KMS는 1개의 노드에서 약 2.3개 키워드를 찾았고, 진화하는 P2P-SN-KMS는 1개의 노드에서 약 7.1개 키워드를 찾은 것이다. 따라서, 진화기능은 P2P-SN-KMS의 검색효율성을 높일 수 있는 중요 기능이며, 자신의 관심그룹 형성도 진화기능이 있는 P2P-SN-KMS이 월등하다는 것을 실험을 통해 확인 하였다.

6. 결론 및 향후 연구과제

지식이 광범위하게 흩어져 있으면서 역동적으로 변화하는 환경에서 지식을 효율적으로 관리하기 위해서는 P2P 기술이 저비용과 편리성을 제공하는 적절한 대안으로 평가 되고 있다. 그러나 P2P 기술은 네트워크를 구성하는 모든 노드를 검색하는 Flooding 방식으로 검색의 효율성이 높지 않은 단점이 있었다. 본 논문에서는 관련성이 높은 지식을 보유한 노드가 물리적으로 멀리 떨어져 있더라도 논리적으로는 바로 이웃에 있도록 하여 지식관리의 효율성을 높일 수 있는 사회적연결망 기법을 도입하여 P2P 환경에서 사회적 연결망을 활용한 지식 관리시스템을 제안 하였다. 논리적으로 가까운(즉, 보유 지식의 관련성이 높은) 이웃노드가 역동적으로 재구성 됨으로써 지식 네트워크가 지식 검색의 결과를 즉각적으로 반영하는 진화의 메커니즘을 구현 하였다. P2P-SN-KMS 시스템의 노드 정보구성, 네트워크 형성, 네트워크 진화에 사용된 다양한 파라미터를 도입하여 시뮬레이션 과정에서 활용함으로써 도메인 지식의 구조에 따른 최적의 파라미터를 찾을 수 있다. 앞으로 예제 문제가 아닌 대량의 노드를 갖는 네트워크 상에서의 검증과 키워드 간의 유사성 평가에 있어서 스트링 매칭이 아닌 온톨로지 기반의 평가를 도입하여 지식의 탐색 및 표현하는 방법에 있어 현재의 시스템보다 더 역동적이고, 네트워크 변화에 능동적으로 대처 할 수 있는 발전된 시스템에 관한 연구가 필요하다.

참고문헌

[1] 김병룡, 김기창, "Gnutella Protocol을 기반한 P2P Web Service 개발", *한국정보과학회 학술발표논문집 2001년도 가을*, (2001), 163~165.

[2] 김영진, 엄영익, "순수 P2P 네트워크 환경을 위한 효율적인 피어 연결 기법", *정보과학회 논문지*, Vol.31, No.1(2004), 11~19.

[3] 김인숙, 강용혁, 엄영익, "순수 P2P 환경에서 분산된 위치 정보를 이용한 자원 검색 기법", *정보과학회논문지: 정보통신*, Vol.31, No.6 (2004), 573~581.

[4] 서치종, 서의호, "Knowledge Network 구축 수단으로서의 Social Network Analysis에 관한 연구", *한국경영과학회 학술대회논문집 학술대회 논문집 2002년 추계*, (2002) 317~322.

[5] 오세현, "P2P 방식의 파일 공유 서비스 : 현황과 전망", *KISDI IT FOCUS*, Vol.2000, No.11 (2000), 3~20.

[6] 오정숙, "KMS(Knowledge Management System)", *정보통신산업동향*, Vol.2, No.2(2001), 51~63.

[7] 이광현, 전형수, 유철중, 장옥배, "P2P 기반 자료 공유 서비스 시스템의 설계 및 구현", *한국정보과학회 학술발표논문집 2002년도 봄*, (2002), 547~549.

[8] 이태식, 이동욱, 배건, "지식관리시스템 구축 방법론에 대한 연구", *2001학술발표회 논문집*, (2001), 810~1813.

[9] Churchill E. F. and C. A. Halverson, "Social Networks and Social Networking", *IEEE Internet Computing*, September/October 2005, 14~19.

[10] Dictionary of e-business 한국전자거래진흥원, (2004).

[11] Fang Wang, Yamir Moreno and Yaoru Sun, "Structure of N-to-N social networks", *American Institute of Physics*, E 73(2006), 036123-1~036123-7.

[12] Ronald Maier and Thomas Hädrich. "Centralized versus N-to-N knowledge management systems", *Journal of Knowledge and Process Management,* Vol.13(2006), 47~61.

[13] Tsoumakos, D. and Roussopoulos, N.. "Adaptive probabilistic search for N-to-N networks", *Proceedings. Third International Conference on*, (2003), 102~109.

[14] Upadrashta, Y. and Vassileva, J. and Grassmann, W. "Social Networks in N-to-N Systems", *Proceedings of the 38th Annual Hawaii International Conference on*, (2005) 200c~200c.

Abstract

Social Network-Based Knowledge Management System for P2P Environment

Youn-Sang Kim* · Suhn Beom Kwon**

P2P (Peer to Peer) techniques have been well applied to file sharing due to its cost-effectiveness and convenience. Dynamic network evolution is another good thing for P2P according to addition and deletion of nodes and change of files a node has. Our research proposes a P2P-based KMS (Knowledge Management System). Knowledge of enterprises spreads all over sub-organizations like oversea factories and sales departments and is changed in dynamic manner. P2P techniques are, therefore well matched with knowledge management domain.

In order to increase search efficiency, we introduce social network theory into P2P-based KMS. Social network technique makes the most similar nodes (in KMS domain, nodes which has the most similar knowledge) its own neighbors, which makes eventually search efficiency increase. We developed our prototype system P2P-SN-KMS and evaluated by simulation.

Key words : P2P, KMS, Social Network

* Nethru Inc.
** School of Business Administration, Kookmin University