

# 실행공동체를 위한 지식관리시스템에서의 퍼지기반 신뢰도 측정

양근우<sup>a</sup>, 허순영<sup>b</sup>

<sup>a</sup> 한국과학기술원 테크노경영대학원  
130-012, 서울특별시 동대문구 청량리동 207-43  
Tel: +82-2-958-3650, Fax: +82-2-958-3604, E-mail: simpact@kgsm.kaist.ac.kr

<sup>b</sup> 한국과학기술원 테크노경영대학원  
130-012, 서울특별시 동대문구 청량리동 207-43  
Tel: +82-2-958-3626, Fax: +82-2-958-3604, E-mail: syhuh@kgsm.kaist.ac.kr

## Abstract

*The importance of communities of practice (CoP) as an organizational informal unit for fostering knowledge transfer and sharing gains a lot of attention from KM researchers and practitioners. Since most of CoPs are formulated online these days, the credibility or trustworthiness of knowledge contents circulated within a certain CoP should be considered thoroughly for them to be fully utilized safely. Here comes the need for an appropriate trust measuring methodology to determine the true value of knowledge given by unknown people through an online channel. In this paper, an improved trust measuring method is proposed using new trust variables such as level of degrees derived from the relationship among community users. In addition, activeness, relevance, and usefulness of the knowledge contents themselves, which are calculated automatically using a text categorization technique, are also used for trust measurement. The proposed framework incorporates fuzzy set and calculation concepts to help build trust matrices and models, which are used to measure the level of trust involved in specific knowledge concerned.*

## Keywords:

CoP; Knowledge Management; Trust Measure

## 서론

오늘날의 치열한 경쟁 환경 하에서 기업의 경쟁력을 유지하기 위해서는 조직 내에 존재하는 유무형의 지식 자산에 대한 관리가 매우 중요한다. 이러한 기업의 요구를 만족시키기 위해 연구되어 온 분야가 지식 관리 (Knowledge Management)라 할 수 있다 [1]. 특히 코드화 또는 시스템화가 상대적으로 어려운 암묵지 (tacit knowledge)의 경우 이를 의도적으로 소유자로부터 분리하기 보다는 필요한 지식을 보유하고 있는 조직내의 전문가를 검색하고 이들과의 의사 소통을 돕는 방향으로 지식 관리 시스템 (Knowledge Management System: KMS)의 기능이 발전하고 있다 [4, 12].

이에 공통의 목표달성을 위해 자발적으로 경험 및 정보를 공유하고 교환하는 비공식적 모임의 의미하는 실행공동체 [13]가 조직의 지식을 전파하고 공유하는 중요한 매개체로서 크게 주목 받고 있다. 오늘날 대부분의 실행공동체는 온라인상의 커뮤니티 형태로 운영되고 있는데 따라서 실제로 대면 접촉이 없었던 실행공동체 회원간에도 상호 정보 교환이 이루어지고 있는 실정이다. 이에 실행공동체를 통해 제공되고 있는 정보 혹은 질의응답 내용의 신뢰도에 관한 고려가 필수적이라 할 수 있다.

본 연구에서는 실행공동체 내의 회원 상호간 인맥 관계, 특정 전문 분야에 대한 활동성 (activeness), 관련성 (relevance), 유용성 (usefulness) 등의 다양한 신뢰 측정 변수를 고려한 신뢰 측정 방법론을 제안하고자 한다. 본 연구에서 제안하는 실행공동체를 위한 신뢰 측정 방법론에서는 다양한 신뢰 측정 변수를 이용하여 특정 회원간의 상대적 상호 신뢰도 및 특정 지식물에 대한 신뢰도를 계산하여 제공하는데 이를 위해 퍼지 신뢰 평면 [8]과 이를 이용한 퍼지값 합성 기법을 도입하여 활용한다.

## 관련연구

지식 공유와 전파의 목적으로 자발적으로 조직되는 실행공동체의 특성상 공동체 회원들간의 참여가 자율적이고 이에 대한 관리나 제약이 없으므로 지식 공유를 위한 편리한 방법을 제공할 수 있는 온라인 커뮤니티 형태의 구성이 매우 일반적이다. 또한 이러한 커뮤니티 형태의 온라인 실행공동체는 시공간적 제약을 극복할 수 있다는 장점도 가진다. 최근에는 불특정 다수의 온라인 커뮤니티 사이트 회원들을 대상으로 하는 지식 장터 개념의 서비스들도 인기를 끌고 있는데 네이버의 지식iN (kin.naver.com), 구글의 Answers (answers.google.com), 엠파스의 지식거래소 (kdaq.empas.com) 등이 그 예라 할 수 있다. 이와 같이 익명성이 보장되는 온라인상의 지식 교류에 있어서 해당 지식물 혹은 제공자의

신뢰도에 대한 고려가 필수적이라 할 수 있는데 이는 잘못된 지식 혹은 정보의 수용으로 인한 피해는 예상보다 훨씬 클 수 있기 때문이다.

온라인상의 신뢰도와 관련한 연구는 주로 전자상거래에서 판매자와 구매자간의 신뢰도 측정 등에 활용되기 위한 방법으로 많이 연구되어 왔으며 [5, 6, 9, 14] 신뢰도 측정을 위해 사용되는 변수 혹은 변수들의 조합을 적절히 선택하는 것이 신뢰도를 정확히 파악하는데 가장 중요한 요소라 할 수 있다. 따라서, 본 연구에서는 새로운 신뢰 측정 변수의 도입을 위해 인터넷 상의 인맥형 커뮤니티 서비스를 활용하였는데 인맥형 커뮤니티란 사용자들간에 현실 세계에서의 실제 관계를 온라인 상에서 구축할 수 있도록 하는 방법을 제공해 주는 서비스로 커뮤니티 내의 사용자 관계 정보 등은 현실 세계의 실제 정보를 나타낸다는 의미에서 신뢰 측정 변수로서의 가치가 높다고 볼 수 있다.

기존의 커뮤니티 서비스와 인맥형 인터넷 커뮤니티를 구분하는 가장 특징적인 요소는 커뮤니티 내의 회원들이 실생활에서의 관계를 바탕으로 서로 연결되어 네트워크를 형성할 수 있다는 사실이다. 예를 들어, 실생활에서 친구관계인 두 사람은 인맥형 커뮤니티 내에서 실제로 친구관계로 연결될 수 있는데 이러한 직접적인 연결관계를 '1촌 관계'라 한다. 또한 2촌 관계의 사람은 한 사람을 중심으로 각각 1촌 관계를 맺고 있는 사람들간의 관계라 할 수 있다. 이러한 관점에서 커뮤니티 내에서 실제 관계의 확장 정도인 촌수는 해당 관계로 맺어진 두 사람간의 친밀도를 나타낸다고 볼 수 있다.

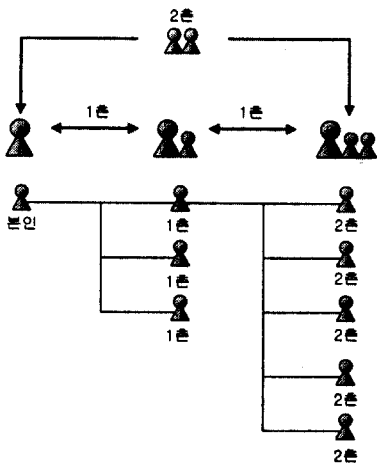


그림 1- 인맥형 커뮤니티내의 인맥연결

인맥형 인터넷 커뮤니티는 회원들이 강제가 아닌 자발적으로 정확한 신상 정보를 입력하도록 유도함으로써 익명성과 관련한 문제를 어느 정도 극복하고 있다. 예를 들어, frenzit.com이나 cyworld.com과 같은 커뮤니티 서비스에서는 자신이 실제로 알고 있는 사람과 서로 원할 경우 '1촌

관계 (first-degree relationship)' 로 연결된다. 이와 같은 원리로 자신과 1촌 관계로 연결된 사람이 다시 누군가와 1촌 관계를 갖는다면 자신은 그 사람과 2촌 관계를 통해 연결된다. 그림 1은 이와 같은 인맥형 커뮤니티 내에서의 회원간 인맥연결의 개념을 나타내고 있다.

인터넷과 관련한 대부분의 문제는 인터넷이 가지는 익명성이라는 특성에서 기인한다고 볼 수 있는데 이는 악의적인 의도를 가진 인터넷 사용자가 자신의 실제 신원이 공개되어 있을 경우 악의적인 행위를 저지를 가능성이 매우 낮기 때문이다. 이러한 의미에서 인맥형 커뮤니티 내에서 1촌 관계를 맺고 있는 사람의 수, 해당 커뮤니티의 평균 방문 회수 등의 회원 프로파일 정보는 그 커뮤니티 내에서의 해당 회원의 충성 정도를 나타낸다는 의미에서 신뢰도를 측정하는데 사용할 수 있는 신뢰 변수로서의 가능성이 높다고 볼 수 있다. 인맥형 커뮤니티 내의 회원들은 커뮤니티 내에서의 네트워크 효과를 극대화하기 위해 자발적으로 자신들의 올바른 신상정보를 입력하는 경향이 있으며 따라서 그들의 신원이 공개되어 있는 공간에서의 악의적인 행위에 관련되는 것을 피하려 하게 된다. 또한 친밀도가 높은 회원들간의 관계에 있어서는 악의적인 행위가 개입될 가능성이 그만큼 낮게 된다.

## 신뢰도 측정을 위한 신뢰 변수

이 장에서는 실행공동체내의 회원간 혹은 특정 등록 지식물에 대한 신뢰도 측정을 위해 도입된 새로운 신뢰 변수들을 소개한다. 회원간의 신뢰도와 특정 지식물에 대한 신뢰도는 그 특성에 따라 적합한 신뢰 변수의 조합을 이용하여 계산하게 된다.

### 회원간의 관계기반 변수

인맥형 인터넷 커뮤니티에서 도입 가능한 추가적인 신뢰 변수가 다수 존재하는데 그 중 하나는 특정 회원간의 촌수이다. 예를 들어, 커뮤니티 내의 두 회원이 2촌의 관계를 가지고 있다면, 이는 그들이 알고 있는 동일인이 해당 회원들과 각각 1촌의 관계를 가지고 있다는 의미가 된다. 따라서, 이런 경우 해당 실행공동체 내에서 두 회원이 정보 제공자와 사용자의 관계를 가질 경우 그들이 전혀 관계를 가지고 있지 않은 사람과의 관계와 비교할 때 더욱 신뢰도가 높은 정보 교환이 이루어질 가능성이 높다고 볼 수 있다.

네트워크 이론에 따르면 상당수의 회원을 가진 커뮤니티 사이트의 경우에도 이러한 인맥관계를 확인해 보면 생각보다 훨씬 적은 촌수 내에서 모든 회원들간의 관계를 정의할 수 있음을 알 수 있다 [3]. 네트워크 이론의 식 (1)은 커뮤니티 회원간의 평균 촌수를 계산하는데 이용할 수 있는 공식으로

전체  $N$ 명의 회원들이 평균적으로  $k$ 개의 1촌 관계를 가진다면 전체 회원들을 연결하는데  $d$ 의 링크 즉 촌수가 필요하다는 것으로 해석될 수 있다. 또한 표 1은 이러한 커뮤니티 회원간의 평균거리 계산의 예를 보여주고 있는데 예를 들어 총 50만명의 회원이 가진 커뮤니티의 경우 각 회원들이 평균적으로 5명의 1촌 관계를 보유하고 있다면 모든 회원간의 관계를 약 8촌 이내에서 정의할 수 있음을 알 수 있다.

$$d = \frac{\log N}{\log k} \quad (1)$$

표 1- 회원간의 평균거리

회원수 \ 평균 1촌수	2	3	4	5
100,000	16.6	10.5	8.3	7.2
500,000	18.9	11.9	9.5	8.2
1,000,000	19.9	12.6	10.0	8.6
5,000,000	22.3	14.0	11.1	9.6
10,000,000	23.3	14.7	11.6	10.0

사용 가능한 또 다른 신뢰 변수로는 해당 회원이 커뮤니티에서 보유한 1촌 관계의 수가 있다. 예를 들어, 어떤 회원이 커뮤니티 내에서 50명의 다른 회원과 1촌 관계를 가지고 있다면 해당 회원의 신원을 확인해 줄 50명의 다른 회원이 그 커뮤니티 사이트에 존재한다는 것과 같은 의미이다. 따라서 어떤 사람이 더 많은 사람과 1촌 관계를 형성하고 있다면 그 사람의 신뢰도는 그렇지 않은 경우와 비교하여 상대적으로 높다고 생각할 수 있다. 일정 기간 동안의 로그인 회수도 또한 그 사람의 해당 커뮤니티 사이트에 대한 충성도를 의미한다는 점에서 사용 가능한 신뢰 변수의 하나로 볼 수 있다. 예를 들어, 하루에 10회 가량 해당 커뮤니티를 찾는 사람은 일주일에 평균 1회 정도 그 커뮤니티를 방문하는 사용자에게 비해 높은 사이트 충성도를 가진다고 볼 수 있다. 또한 이러한 높은 충성도는 결과적으로 높은 신뢰도를 의미하는 것으로 해석이 가능하다.

### 전문지식 관련 변수

실행공동체에서 제공되는 지식물 또는 소속 전문가가 보유하고 있는 전문 지식은 활동성, 관련성, 유용성의 세 가지 요소를 기준으로 측정될 수 있으며 각 요소에 대한 개념과 측정 방법은 다음과 같다.

#### 활동성 (activeness)

본 연구에서 활동성은 실행공동체내의 전문가가 커뮤니티에 지식물을 등록하는 지식 생성 활동을

얼마나 빈번히 수행하는지를 의미한다. 활동성은 또한 빈도 (frequency), 최신도 (recency), 분량 (volume)의 요소로 구성되어 있다고 볼 수 있다. 빈도란 특정 전문가에 의해 시스템에 등록된 지식물의 숫자를 의미한다. 더 많은 지식을 등록한 전문가가 해당 분야에 대한 전문 지식을 상대적으로 더 많이 보유하고 있다고 볼 수 있는 것이다. 지식은 시간이 지남에 따라 그 가치가 변하게 되는데 최신도는 이러한 지식의 특성을 지식물의 평가에 반영하도록 한다. 따라서, 지식물은 그 가치를 평가함에 있어 시간 요소에 의해 조정되어야 할 필요가 있는 것이다. 마지막으로 분량 요소는 양이 방대한 지식을 등록한 전문가에 대한 고려를 위해 도입된다. 일반적으로 말해서, 분량이 큰 지식물이 더욱 많은 정보를 전달할 수 있다고 볼 수 있는데 따라서 시스템에서는 이러한 분량이 큰 지식물을 등록하는데 드는 시간과 노력을 인정해 줄 필요가 있는 것이다. 아래의 식 (2)은 이 절에서 설명한 전문가의 활동성을 수식으로 표현하고 있다.

$$\sum_{i=1}^n w_{T_i}(T_i) * w_{V_j}(V_j) \quad (2)$$

식 (2)에서  $T_i$ 와  $V_j$ 는 지식물  $i$ 가 등록된 시간과 그 분량 요소를 각각 의미한다. 또한,  $w_{T_j}$ 는 주제 분야  $j$ 에서의 최신도에 대한 가중치를 의미하며 마찬가지로  $w_{V_j}$ 는 분야  $j$ 에서의 분량 가중치를 의미한다. 특정 주제 분야를 가지는 실행공동체는 등록되어 있는 전체 지식물의 내용이 커뮤니티에 따라 서로 다르게 되며 따라서 새롭게 등록되는 지식물에 대한 최신도와 분량 요소를 결정하기 위해 사용되는 기준이 서로 다르게 적용 된다. 이에 이러한 주제 분야간의 차이를 고려한 최신도와 분량에 대한 가중치를 적용하여 전문 분야 지식 측정치를 수정해 줄 필요가 있다. 예를 들어,  $T_i$ 의 경우 특정 기준 시점으로부터 해당 지식물이 등록된 시점까지의 일수 혹은 경과 시간 등을 이용할 수 있으며  $V_j$ 는 해당 지식물의 단어 혹은 글자수 등을 활용할 수 있다. 또한 이들에 대한 가중치인  $w_{T_j}$ 와  $w_{V_j}$ 는 해당 분야에 따라 달라지게 되는데 이는 지식물이 등록되는 빈도와 평균적인 지식물의 분량 등이 분야마다 다른 양상을 보이기 때문이다.

한편, 빈도 요소의 경우 다른 요소의 취합을 통해 자동적으로 반영되는 것으로 볼 수 있는데 이는 모든 지식물이 이러한 활동성 계산에 의해 영향을 받기 때문이다. 즉, 전문가가 더 많은 지식물을 등록할수록 해당 전문 분야에 대한 전문 지식을 해당 전문가가 더 많이 보유하고 있다고 생각할 수 있다. 특정 전문가에 대해서 식 (2)을 이용해 계산된 모든 값의 합은 특정 실행공동체의 전문 분야에 대한 해당 전문가의 활동성 정도를 나타낸다.

**관련성 (relevance)**

전문 지식을 구성하는 두 번째 요소인 관련성은 등록된 지식물이 얼마나 특정 주제 분야에 적합한가를 의미한다. 등록된 지식물이 속하는 주제 분야와 해당 분야에의 적합도를 계산하는 것이 가능하다면 우리는 이러한 요소를 이용하여 해당 지식물이 특정 주제 분야와 얼마나 관련이 있는가를 결정할 수 있게 된다. 관련성은 다음과 같은 식을 이용하여 계산될 수 있다.

$$\sum_{j=1}^n S_j(D_i) \quad (3)$$

식 (3)에서  $D_i$ 는 문서 혹은 지식물  $i$ 를  $S_j$ 는 해당 지식물의 주제 분야  $j$ 에 대한 관련성 점수를 각각 의미한다. 관련성 점수  $S_j$ 는 학습용 문서를 이용하여 미리 학습 과정을 거친 지식 자동 분류기에 의해 측정되는데 다음과 같은 과정을 거친다. 우선, 특정 전문가에 의해 지식물이 커뮤니티에 등록되면 (단계 1) 지식 자동 분류기가 호출되어 새로 등록된 지식물을 분석하게 된다. 우선 분류기는 지식물로부터 문서 벡터를 생성하게 되며 (단계 2) 등록된 지식물로부터 생성된 문서 벡터는 문서 대 분야의 쌍대 비교를 통해 각 주제 분야와의 유사도를 측정하는데 이용된다 (단계 3). 이 과정에서 벡터 유사도 계산법 [2]을 도입하여 활용하게 되며 이전 단계의 결과로 선택된 주제 분야와 해당 주제 분야에 대한 유사도는 해당 지식물의 분야 관련성 점수가 되고 또한 해당 전문가의 프로필 정보를 갱신하는데 이용된다 (단계 4). 그림 3는 이와 같이 지식 자동 분류기를 이용하여 등록된 지식물을 자동으로 분류하는 과정을 도식화하고 있다.

중요한 요소 중 하나로 볼 수 있다. 실제로 매우 가치 있는 지식이 커뮤니티에 등록되었다 하더라도 커뮤니티내의 다른 사용자가 해당 지식물을 이용하고 그로부터 해당 지식의 가치를 인정하기 전에는 그 지식이 유용하거나 의미 있는 것이라고 보기 힘들다. 따라서, 등록된 지식에 대한 다른 사용자의 평가를 전문 지식을 측정하는 과정에 포함시키는 것이 필요하다. 평가 요소는 다음과 같이 표현될 수 있다.

$$\sum_{i=1}^n \sum_{k=1}^m A_k(D_i) \quad (4)$$

식 (4)에서  $A_k$ 는 사용자  $k$ 에 의해 평가된 지식물  $i$ 에 대한 평가 점수를 나타낸다. 하나의 등록된 지식물은 한 명 이상의 사용자에게 의해 활용되고 평가될 수 있으므로 지식물  $i$  ( $D_i$ )의 사용자로부터 받은 모든 평가 점수를 합산하여 해당 지식물의 평가 점수를 도출하게 된다. 또한 특정 전문가가 등록한 모든 지식물에 대한 평가 점수를 합산하면 해당 전문가의 전체 평가 점수가 되는데 식 (4)은 이것을 표현하고 있다.

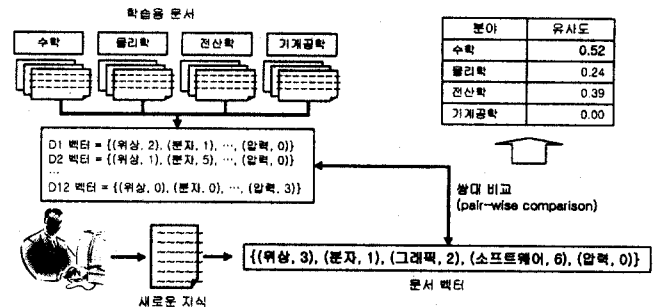


그림 3- 지식 자동 분류

**유용성 (usefulness)**

유용성 역시 등록된 지식물의 가치를 평가하는데

**CoP를 위한 신뢰 측정 시스템의 설계**

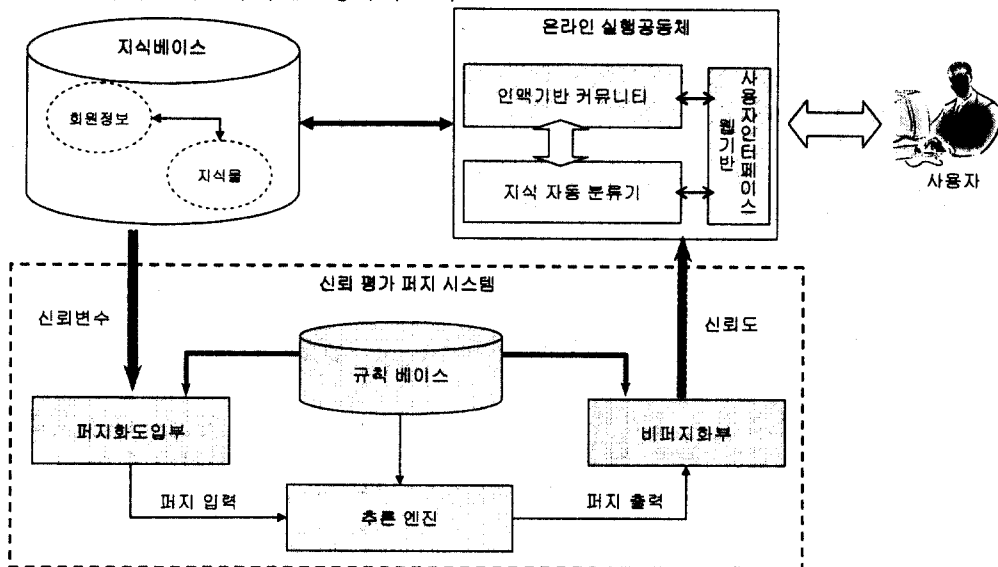


그림 2- 실행공동체를 위한 신뢰 측정 시스템

이 장에서는 본 연구에서 제안하는 다양한 신뢰 측정 변수들을 이용하여 최종 신뢰도를 계산하는 방법을 소개한다. 신뢰도 계산을 위해 모호한 값의 처리가 가능한 퍼지 이론 [15]을 도입하였으며 이를 통해 다양한 신뢰 변수로부터 도출된 출력 값들은 하나의 최종 신뢰도 값으로 변환된다.

### 신뢰 측정 시스템 아키텍처

본 연구에서 제안하는 신뢰 측정 시스템의 구조는 크게 두 개의 하부 구성요소로 구분될 수 있다. 그 하나는 실행공동체 회원들간의 지식 공유와 전파를 위한 기능을 제공하는 인맥형 커뮤니티 서비스로 회원들은 다른 온라인 커뮤니티를 이용하는 것과 동일한 방식으로 자신이 가입한 실행공동체 커뮤니티를 통해 지식 공유를 할 수 있다. 다른 한 부분은 이러한 커뮤니티 서비스와 연계된 신뢰 평가 퍼지 시스템으로 커뮤니티 내에서 제공되는 지식물에 대해 실시간으로 신뢰도를 측정하는 것을 담당한다. 그림 2는 제안된 시스템의 전체적인 구조를 보여 주고 있다. 이 중 신뢰 평가 퍼지 시스템의 목적은 지식 공유에 참여하는 커뮤니티 회원간의 신뢰도를 정량화하는 것으로 퍼지화 도입부, 추론 엔진, 비퍼지화부 등의 하부 구조를 가지고 있다. 그림 2에서 규칙 베이스는 신뢰도를 측정하기 위해 추론 시 사용되는 퍼지 규칙을 정의해 놓은 부분이다.

지식 공유에 참여하는 회원간 혹은 등록된 지식물의 신뢰도를 측정하기 위해 제안된 신뢰 측정 시스템은 각 회원의 분야별 전문 지식의 수준은 물론 인맥형 커뮤니티 서비스에서 도입한 사용자 프로파일 정보 등도 새로운 신뢰 측정 변수로 이용한다. 제안된 실행공동체 커뮤니티의 세부 구조에 대한 자세한 설명은 다음과 같다.

### 신뢰도 측정

앞 장에서 소개한 신뢰 변수들 간의 관계는 퍼지 신뢰 평면 (Fuzzy Trust Surface) [6]이라 불리는 매트릭스를 이용하여 표현할 수 있다. 그림 4의 (a)는 특정 지식물의 특정 분야에 대한 관련성이라는 신뢰 변수와 최신도라는 신뢰 변수 사이의 관계를 표현하는 퍼지 신뢰 평면 매트릭스의 예를 보여주고 있다. 매트릭스 내부 각 셀의 내용은 각 신뢰 변수에 대한 입력 값의 조합에 따라 해당 거래에 부여되는 신뢰도를 나타낸다. 한편 그림 4의 (b)는 본 연구에서 새롭게 도입한 신뢰 변수 중 하나인 관련 회원간 촌수와 전문지식 수준의 관계를 보여주고 있다. 이러한 퍼지 신뢰 평면 매트릭스와 그 셀값은 해당 분야의 전문가에 의해 정의될 수 있는데 이렇게 정의된 퍼지 매트릭스는 각 지식물 혹은 전문가에 대한 최종 신뢰도를 계산하기 위해 퍼지 전문가 시스템의 퍼지 추론 엔진에서 사용되는

퍼지 규칙으로 변환된다.

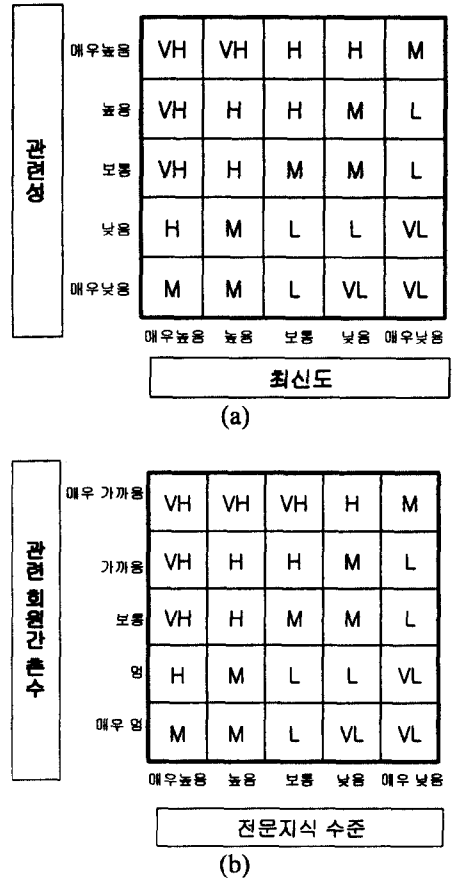


그림 4- 퍼지 신뢰 평면

### 퍼지화 도입부

퍼지화 도입부는 입력 값에 대한 퍼지화를 담당하는 부분으로 각 입력 값을 그 종류에 따라 미리 정의된 퍼지 값 중의 하나인 언어적 표현으로 바꾸는 역할을 담당한다.

앞 장에서 설명된 신뢰 변수들을 퍼지 추론 엔진을 통해 신뢰도 평가에 사용하기 위해서는 먼저 이러한 신뢰 변수들에 대한 입력 값을 퍼지 수 (fuzzy number)의 형태로 변환할 필요가 있다. 이러한 변환은 신뢰 평가 퍼지 시스템의 퍼지화 도입부에 의해 수행된다. 또한 규칙 베이스에는 각 신뢰 변수에 대한 퍼지 소속 함수가 미리 정의되어 있다. 사다리꼴 퍼지 함수를 사용하여 거래 개체간의 촌수에 대한 소속 함수를 다음의 예와 같이 정의해 볼 수 있으며 그림 5는 이러한 소속 함수 정의를 그림으로 보여 주고 있다.

- 매우 가까움 = [1, 1, 1, 4]
- 가까움 = [1, 3, 4, 6]
- 보통 = [3, 5, 6, 8]
- 멀 = [5, 7, 8, 10]
- 매우 멀 = [7, 9, 15, 15]

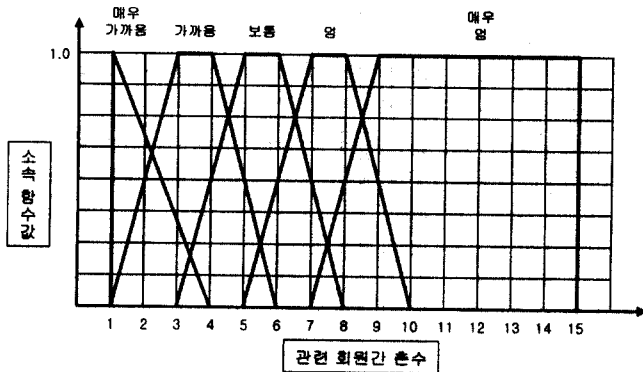


그림 5- 퍼지 소속 함수의 예

예를 들어, '매우 가까움'이라는 원소에 대해서는 1촌에서 시작하여 1촌, 또 다시 1촌을 연결하고 마지막으로 4촌을 연결하는 형태의 소속함수를 가진다는 의미로 이 경우에는 최대값을 가지는 두 점이 1촌으로 동일하여 삼각형 형태의 소속함수를 가지게 된다. 다음으로 '가까움'이라는 원소에 대해서는 1촌을 시작으로 3촌과 4촌 사이에서 최대값을 가지고 다시 6촌을 최소값으로 연결하는 사다리꼴 형태의 소속함수를 가진다는 의미이다.

이와 같이 각 신뢰 변수에 대해 정의된 퍼지 소속 함수를 이용한 명확한 형태의 입력 값은 퍼지 수 형태의 값을 가지도록 변환 과정을 거치게 된다. 우선, 입력 값들은 퍼지합이나 퍼지곱 등의 퍼지 연산을 수행하는데 편리하도록 미리 정의된 입력 퍼지 값의 허용 범위에 맞도록 크기 변환 (scale mapping) 과정을 거치게 된다. 다시 이렇게 변환된 값들은 적합도 (DOF: Degree of Fulfillment)를 계산하는 과정을 거쳐서 퍼지화 된다. 적합도란 해당 입력 값이 고려하는 퍼지 소속 함수에 대해 어느 정도의 수준으로 대응되는 지를 나타내는 값으로 예를 들어 관련 회원간의 촌수라는 신뢰 변수에 대해 입력 값이 4라면, 해당 퍼지 집합의 '매우 가까움' 원소에 대한 퍼지 소속함수 값은 '0'이 되고 '가까움'이라는 원소에 대한 소속함수 값은 '1'이 된다. 이러한 방식으로 관련 회원에 대한 전체 적합도를 구해 보면

$$DOF_{\text{관련 회원간 촌수}=4} = \{0, 1, 0.5, 0, 0\} \text{ 과 같다.}$$

이렇게 계산된 적합도는 퍼지 추론 엔진에서 규칙 베이스에 정의된 신뢰 변수 간의 관계에 따라 두 변수간의 적합도 매트릭스를 구성하는데 사용되며 이는 다시 최종 신뢰도 계산을 위해 매트릭스 합성 단계를 거치게 된다.

### 퍼지 추론 엔진

퍼지 추론 엔진은 실제로 퍼지 논리와 규칙 베이스에 정의된 퍼지 규칙을 이용하여 신뢰도 측정을 담당하는 부분이며 인간의 의사 결정 과정과 유사한 방식으로 퍼지 논리와 퍼지 규칙을 기반으로 신뢰도를 계산하게 된다.

퍼지화 도입부에 의해 퍼지 수로 변환된 입력 값들은 퍼지 추론 엔진에 의해 신뢰도를 계산하는데 사용된다. 규칙 베이스에는 시스템 내의 모든 입력 변수에 대한 퍼지 소속 함수와 각 변수들간의 관계가 미리 정의되어 있다. 규칙 베이스 내에 정의되어 있는 신뢰 변수간의 관계에 따라 적합도 계산을 위해 두 신뢰 변수가 선택되고 다음의 식 1에 따라 두 변수 간의 적합도 합성 결과 매트릭스에 최소 값이 부여된다.

$$w_i = A_{i1}(x_1^0) \wedge A_{i2}(x_2^0) \quad (5)$$

$$C_0(w) = [w_1 \wedge C_1(w)] \vee [w_2 \wedge C_2(w)] \vee \dots \vee [w_n \wedge C_n(w)] \quad (6)$$

$$= \max[w_i \wedge C_i(w)] \quad \text{for } i=1 \text{ to } n$$

식 (5)에서  $X_1^0$ 와  $X_2^0$ 는 입력 값을 나타내며  $A_{i1}$ 과  $A_{i2}$ 는 각각 해당 신뢰 변수의 퍼지 소속 함수를 나타낸다. 또한 '∧'는 최소 연산자 (min operator)로 퍼연산자 중 더 작은 값을 결과로 돌려준다. 따라서 최대-최소 연산 (Max-Min operation) [15]을 이용할 경우, 최종 신뢰도인 식 (6)의  $C_0(w)$ 는 다양한 신뢰 변수들의 쌍으로부터 계산된 퍼지 출력 값들 중 최대 값에 의해 결정된다. 예를 들어, 관련 회원간의 촌수와 전문지식 수준이라는 두 변수에 대해서 입력 값이 각각 4와 125라고 하자. 두 입력 값에 대한 적합도는 다음과 같이 주어진다.

$$DOF_{\text{관련 회원간 촌수}=4} = \{0, 1, 0.5, 0, 0\}$$

$$DOF_{\text{전문지식 수준}=125} = \{0, 0.25, 1, 0.25, 0\}$$

주어진 두 입력 값에 대한 각각의 적합도를 합성한 두 변수 간의 최종 적합도는 관련 회원간 촌수를 세로축으로 하고 전문지식 수준을 가로축으로 하여 교차하는 적합도 값 중 더 작은 값을 각 셀의 값으로 선택하며 이렇게 구성된 적합도 매트릭스와 그림 4의 (b)에 제시된 두 변수에 대해 정의된 퍼지 규칙 매트릭스는 하나의 결과 매트릭스로 합성된다. 적합도 매트릭스에서 '0'을 가지는 셀은 퍼지 규칙 매트릭스와의 합성 시 빈 셀로 나타나게 되고 값을 가지는 셀들은 해당 값과 퍼지 규칙 매트릭스 상의 같은 위치에 해당하는 셀이 가지는 값과 통합되어 표시되어 진다.

이러한 합성의 과정을 시각적으로 표현한 결과는 그림 6과 같다. 그림에서 점선은 출력 변수의 신뢰도에 대한 퍼지 소속 함수를 나타내며 매트릭스 합성의 결과인 0.25M, 0.5M, 0.25H, 1H 등은 실선으로 표시되어 있다. 최대-최소 공식을 이용하여 최대 값을 구하면 그림 6에서 실선으로 표시된 구역의 외곽선 부분을 얻게 된다.

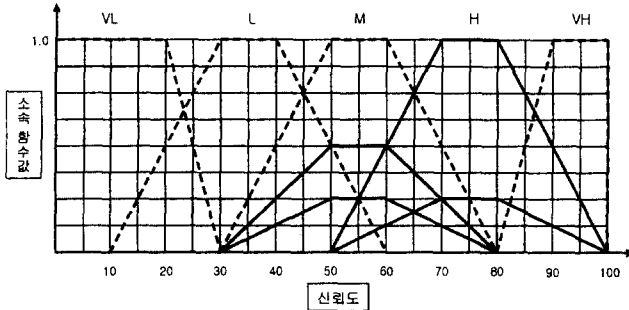


그림 6- 매트릭스 합성 그래프

### 비퍼지화부

퍼지 신뢰 평가 시스템의 비퍼지화부는 추론 엔진에 의해서 계산된 퍼지 출력 값을 정량적인 값으로 변환하는 역할을 한다. 퍼지 추론 엔진은 규칙 베이스에 정의된 신뢰 변수 쌍에 대한 관계의 숫자만큼 결과 값을 출력하게 되므로 고려 중인 지식물 혹은 전문가에 대한 최종 신뢰도를 결정하기 위해서는 다수의 퍼지 출력 값을 하나의 정량화된 숫자로 변환할 필요가 있다. 실제로 이러한 신뢰 평가 시스템의 사용자 입장에서 보면 다양한 퍼지 출력값을 이용 가능한 하나의 신뢰도 점수(measure)로 제공하는 것이 사용 편리성이나 이해도를 높일 수 있는 방법이라 할 수 있다. 제안된 시스템에서는 이러한 변환 과정을 위해 계산이 다소 복잡하지만 주어진 퍼지 집합을 가장 잘 대표하는 값을 생성하는 것으로 알려진 무게 중심법(center of area method)을 사용한다. 즉, 고려하는 퍼지 출력 값의 수만큼 그려지는 매트릭스 합성 결과 그래프들의 전체 면적에서 무게 중심이 되는 지점을 모든 퍼지 결과값을 통해 계산된 최종 신뢰도 값으로 결정하게 된다. 무게 중심법을 이용한 퍼지 결과값 합성의 예로 등록된 지식물의 관련성과 최신도 변수를 통해 계산된 퍼지 결과값과 관련 회원간의 촌수와 전문지식 수준을 이용하여 계산된 퍼지 값과의 합성이 그림 7에 표현되어 있다. 그림 7에서 굵은 실선으로 표시된 지점이 두 퍼지 결과값을 가장 잘 대표하는 최종 신뢰도의 명확한 수치 값을 표시하고 있다.

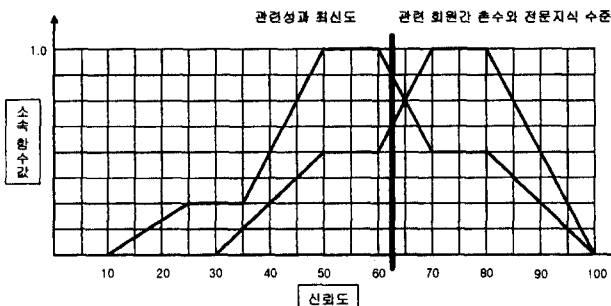


그림 7- 무게 중심법을 이용한 퍼지 결과값 합성

### 요약 및 결론

본 연구에서는 실행공동체에서 회원들간에 지식공유를 수행하거나 해당 실행공동체에 등록된 지식물을 이용하고자 할 때, 상대 전문가 혹은 특정 지식물의 상대적인 신뢰도를 다양한 신뢰 변수로부터 측정된 신뢰 값과 그들의 합성을 통해 제시하는 퍼지 기반 신뢰도 측정 시스템의 구조를 제안하였다. 제안된 시스템에서 사용하는 새로운 신뢰 측정 변수들을 제안하였는데 이 중 각 지식물 혹은 전문가의 전문 지식에 대한 특성을 나타내는 변수들은 자동 문서 범주화 기법을 이용한 지식 자동 분류기에 의해 도출된다. 또한, 지식 공유에 참여하는 관련 회원들간의 관계를 기반으로 하는 정보들도 인맥형 커뮤니티 서비스에서 제공할 수 있는 신뢰 변수들로 활용하였다.

지금까지의 전문가의 전문지식에 관한 프로파일 정보 수집은 수동으로 이루어져 왔는데 이 방식은 (1) 지속적인 관리 비용 요구, (2) 주관적인 판단 기준 혹은 갱신 미비로 인한 오류 발생 가능성 등의 단점을 가지고 있다. 본 연구에서는 이러한 수동 프로파일 정보 수집의 단점을 개선하고자 지식 관리 시스템에 등록된 전문가의 전문 분야 정보를 자동으로 수집하고 특정 주제 분야에 대한 전문 지식의 수준을 측정할 수 있는 지식 자동 분류기의 구조를 제안하였다. 이를 통해 지식물 혹은 전문가의 전문지식의 수준을 도출하여 신뢰도 계산에 활용하였다.

본 연구에서 제시한 퍼지 기반 신뢰 측정 시스템은 온라인 상에서 상대방의 신뢰도 측정이 필요한 다양한 시스템에 적용될 수 있으므로 그 활용 가능성이 높다고 할 수 있다. 향후 연구로는 퍼지 값 계산 과정을 개선하기 위한 추론 엔진의 최적화와 퍼지 규칙의 자동 생성을 위한 방법론 개발 등이 있다.

### Acknowledgments

본 연구는 대학 IT연구센터 육성·지원사업의 연구결과로 수행되었음.

### References

- [1] Alavi, M., and Leidner, D.E. (2001). "Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues," *MIS Quarterly*, Vol. 25, No. 1, pp. 107-136.
- [2] Baeza-Yates, R., and Riberio-Neto, B. (1999). *Modern Information Retrieval*. New York, NY: ACM Press.
- [3] Barabasi, A.L. (2002). *Linked: The New Science of Networks*. Cambridge, MA: Perseus Publishing.
- [4] Desouza, K.C. (2003). "Barriers to Effective Use of Knowledge Management Systems in Software Engineering," *Communications of the ACM*, Vol. 46, No. 1, pp 99-101.

- [5] Doney, P., and Cannon, J. (1997). "An Examination of the Nature of Trust in Buyer-Seller Relationships," *Journal of Marketing*, Vol. 61, pp. 35-51.
- [6] Hoffman, D., Novak, T., and Peralta, M. (1999). "Building Customer Trust Online," *Communications of the ACM*, Vol. 42, No. 4, pp. 80-85.
- [7] Kent, S. (1993). "Internet Privacy Enhanced Mail," *Communications of the ACM*, Vol. 36, No. 8, pp. 48-60.
- [8] Manchala, D. (2000). "E-Commerce Trust Metrics and Models," *IEEE Internet Computing*, Vol. 4, No. 2, pp. 36-44.
- [9] Maurer, U. (1996). "Modeling a Public-Key Infrastructure," *Proceedings of the 1996 European Symposium on Research in Computer Security*.
- [10] Mayer, R., Davis, J., and Schoorman, F. (1995). "An Integrative Model of Organizational Trust," *Academy of Management Review*, Vol. 20, No. 3, pp. 709-734.
- [11] McCallum, A.K. (1996). *Bow: A Toolkit for Statistical Language Modeling, Text Retrieval, Classification and Clustering*. <http://www.cs.cmu.edu/~mccallum/bow>.
- [12] Rus, I., and Lindvall, M. (2002). "Knowledge Management in Software Engineering," *IEEE Software*, Vol. 19, No. 3, pp. 26-38.
- [13] Wenger, E., McDermott, R., and Snyder, W.M. (2002). *Cultivating Communities of Practice*. Boston, MA: Harvard Business School Press.
- [14] Yahalom, R., and Klein, B. (1997). "Trust Relationships in Secure Systems - A Distributed Authentication Perspective," *Proceedings of the 4th ACM Conference on Computer and Communications Security*.
- [15] Zadeh, L. (1973). "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *IEEE Transactions on Systems Management and Cybernetics*, Vol. SMC-3, No. 1, pp. 28-44.