

# 다중 평가 기준을 고려한 정량적 신뢰모델

## A Quantitative Trust Model with consideration of Multiple Evaluation Criteria

이건명\*\*, 김학준\*\*

\*충북대학교 전기전자컴퓨터공학부 첨단정보기술연구센터\*

\*\*호원대학교 멀티미디어정보학과

Keon Myung Lee, Hak Joon Kim

Dept. of EECE and AITrc, Chungbuk National University

Division of Multimedia Information, Howon University

E-mail : kmlee@cbnu.ac.kr

### 요 약

이 논문에서는 개체에 대한 신뢰도를 계산하기 위해 여러 가지의 평가기준을 이용하고, 또한 다른 개체들로 부터의 추천정보를 이용하는 신뢰모델에 대해서 제안한다. 제안한 모델에서는 개체의 신뢰도를 개체가 주어진 상황에서 만족스러운 결과를 낼 기대값으로 정의한다. 다른 개체와 상호작용이 일어날 때마다 각 평가기준에 따른 평가결과가 얻어진다고 전제하는 상황에서 적용되는 신뢰 모델이다. 제안된 모델에서는 신뢰정보가 요구될 때 우선 결과확률 분포(outcome probability distribution)와 개체의 평가결과에 대한 선호도를 고려하여 각 평가기준에 대한 만족정도를 계산한다. 이렇게 계산된 만족정도 값들은 각 평가기준의 중요도를 반영하여 하나의 신뢰값으로 결합된다. 이때 추천 정보도 신뢰값에 함께 결합되는 모델이다.

### 1. 서론

전자상거래, P2P 서비스 등과 같은 인터넷 응용이 광범위하게 사용되어감에 따라, 사용자는 인터넷을 통해서 잘 모르는 사용자나 시스템과 거래를 하면서 얼마간의 위험을 감수할 수 밖에 없는 상황이다. 일상에서는 상대방과의 과거 직접경험에 비추어 보거나 명성(reputation)을 참고하여 상대에 대한 신뢰도를 나름대로 평가하여 거래여부에 반영한다. 같은 맥락에서 온라인의 개체도 상대에 대한 신뢰정보를 이용할 수 있으면 위험부담을 줄일 수 있다. 여러 가지 신뢰모델이 지금까지 제안되어 왔지만, 신뢰자체에 대한 정의가 다양할 수 있기 때문에 어떤 모델도

모두를 만족시키지는 못한 상황이다[1-9]. 모델 중에는 정성적이나 정량적으로 정의된 것도 있고, 어떤 모델은 사용자의 평가값을 기초하여 신뢰도를 계산하고, 다른 모델은 일정기간 개체의 행태를 관찰하여 신뢰값을 계산하기도 한다.

신뢰 자체에 대해서는 여러 가지 정의가 있는데, Gambetta는 신뢰를 “어떤 상대 개체가 자신의 행동을 결정할 수 있는 상황에서 특정 행동을 할 것이라는 것에 대한 주관적인 확률”이라고 정의하고 있다. 신뢰모델에서는 일반적으로 상황 신뢰(situational trust), 성향 신뢰(basic/dispositional trust), 일반신뢰(general trust) 등 세가지 신뢰에 대해서 고려한다. 상황 신뢰는 특정 상황에서 다른 개체에 대해서 어떤 개체가 갖는 신뢰이다. 성향 신뢰는 어떤 개체가 다른 개체에 갖는 무조건인 신뢰이다. 일반 신뢰는 상황에 무관하게 어떤 개체가 특정 개체를 믿는 신

\* 이 연구는 첨단정보기술연구센터(AITrc)를 통해서 한국과학재단으로부터 지원을 받아 수행한 것임.

되이다. 개체에 대한 신뢰는 어떤 관점에서 평가하는가에 따라 달라질 수 있다. 예를 들면, 어떤 식당에 대한 신뢰에 관심이 있다고 하자. 어떤 이는 음식의 맛에 관심을 갖고, 어떤 사람은 기다리는 시간에 신경을 더 쓸 수도 있고, 혹은 자신이 선호하는 음식 메뉴가 항상 있는지에 대해서 관심을 가질 수 있고, 어떤 이는 이러한 모든 것에 관심을 가질 수도 있다. 이 논문에서는 이와 같이 여러 가지 평가기준을 고려하여 개체의 신뢰도를 계산하는 모델을 제안한다.

제안하는 모델에서는 신뢰를 개체가 주어진 상황에서 만족스러운 결과(outcome)을 낼 확률로 정의한다. 여기에서는 개체와 상호작용을 할 때마다, 각 평가기준 관점에서 평가가 이루어진다고 전제한다. 어떤 평가는 관찰에 의해서 이루어지고, 어떤 평가는 사용자가 직접하는 경우도 있다. 제안한 모델에서는 신뢰정보가 필요할 경우, 대상 개체에 대한 각 평가기준별로 만족스러운 결과를 낼 확률인 만족도(satisfaction degree)를 계산한 다음, 이들 만족도값을 사용자의 선호도를 감안하여 하나의 신뢰값을 결합하는 방법으로 신뢰를 계산한다.

논문은 다음과 같이 구성된다. 2절에서는 기존의 신뢰모델에 대해서 간단히 살펴보고, 3절에서는 제안한 신뢰모델에 대해서 소개하고, 4절에서 결론을 맺는다.

## 2. 관련연구

신뢰와 명성의 개념은 경제학, 분산인공지능, 에이전트 등 다양한 분야에서 관심을 끌여온 분야이다. 그 결과로 다양한 모델이 제안되어 왔는데, 어떤 것은 이론적인 가이드라인만 제시한 것도 있고, 어떤 것은 계산모델로 구체화한 것도 있다[1-9].

Abul-Rahman 등은 신뢰등급을 very trustworthy, trustworthy, untrustworthy, very untrustworthy 등 네 단계로 나타내는 정성적인 모델을 제안했다. 이 모델은 신뢰등급 및 가중치를 결정하는데 있어서 다소 원칙이 없다는 것이 약점이다.

Azzedin 등은 P2P 네트워크 컴퓨팅 시스템에 대한 신뢰모델을 제안하였는데, 이는 구조화된 네트워크 컴퓨팅 시스템에 대해서 개발되었기 때문에 일반적인 경우에 적용하는데는 제약이 있다.

Derbas 등은 모바일 에이전트를 악의적인 시스템으로부터 보호하기 위해 사용할 수 있는 명성(reputation)에 기반한 TRUMMAR라는 신뢰모델

을 제안하였다. 이 모델은 신뢰를 명성정보를 이용하여 계산하는데 중점을 두고 있는 것으로, 여러 평가기준을 고려하지는 않고 있다.

Shi 등은 개체의 행동에서 나올 수 있는 결과에 대한 통계적 정보를 이용하는 신뢰모델을 제안하였다. 이 모델에서는 신뢰를 하나의 스칼라 값으로 나타내는 대신에 결과에 대한 확률분포로 나타낸다. 거래 대상을 선택할 때, 후보 대상들의 행동에 대한 기대유용성(expected utility)를 계산하는데, 이때 결과에 대한 확률분포를 이용한다.

Wang 등은 P2P 네트워크에서 사용할 수 있는 Bayesian network 기반의 신뢰모델을 제안하였다. 이러한 Bayesian network은 여러 관점에서 개체를 신뢰할 수 있는 확률을 나타낸다. 이 모델은 다른 개체로부터의 추천 정보를 모델에 반영할 수 있도록 하고 있다.

## 3. 제안한 신뢰모델

신뢰를 구성하는 것이 무엇이고, 어떻게 신뢰를 정의할지에 대해서는 아직 전반적인 동의가 이루어지지 않은 상황이다. 이 논문에서는 계산가능한 신뢰모델을 제안하기 위해, 상황신뢰를 '어떤 개체가 주어진 상황에서 평가항목 관점에서 평가될 때 만족스러운 결과를 낼 기댓값이다'라고 정의한다. 이 절에서는 이러한 상황신뢰에 대한 정의에 기반하여 상황신뢰를 평가하는 방법, 성향신뢰, 일반신뢰를 다루는 방법, 다른 개체로부터의 추천을 이용하는 방법, 추천자에 대한 추천신뢰도를 관리하는 방법에 대해서 기술한다.

### 3.1 상황신뢰

상황신뢰는 특정 상황에서의 개체에 대한 신뢰이기 때문에, 상황신뢰값은 상황에 의한 결정된다. 대부분의 기존 신뢰모델에서는 대상 개체의 행동이 얼마나 만족스러웠는지에 대해서 관심을 갖고, 만족도에 대한 평가값을 이용하는 방식을 채택하고 있다. 이 논문에서는 계산적인 신뢰모델을 제공하기 위해, 여러 평가기준을 고려하고 평가개체의 선호도를 반영하는 방법을 제안한다. 제안한 방법에서는 상황신뢰를 다음과 같이 계산한다. 우선, 평가개체는 대상개체와의 거래결과를 평가기준별, 상황별로 누적하여 확률분포를 만든다. 대상개체에 대한 신뢰정보가 필요할 때는 해당 상황에 대해서 각 평가기준에 대한 만족정도를 계산한 다음, 평가항목별 만족정도를 결합하

여 최종 신뢰값을 구한다.

편의상  $TS_\alpha(\beta, \delta, EC)$ 가 객체  $\alpha$ 가 개체  $\beta$ 에 대해서 상황  $\delta$ 에 평가항목집합  $EC = \{ec_1, ec_2, \dots, ec_n\}$  측면에서 갖는 상황신뢰도라고 하자. 여기에서  $ec_i$ 는 각 평가항목을 말한다. 상황신뢰도는 주어진 상황에 평가항목들 관점에서 개체  $\beta$ 가 만족스러운 행동을 할 것이라는 기대값을 나타낸다. 상황신뢰를 계산하기 위해서는,  $\alpha$ 가  $\beta$ 와 거래를 할 때 마다, 각 평가항목 측면에서의 결과값을 누적하여 확률분포에 반영한다. 평가결과는 연속값 또는 카테고리값으로 주어진다. 연속값인 경우에는 미리지정한 몇 개의 구간의 정량화시켜서, 결과값을 구간에 대응하는 대표값으로 표현한다.

### 3.1.1 결과의 확률분포

$PG^t(\alpha, \beta, \delta, o_i; ec_k)$ 는  $\alpha$ 가 관찰한  $\beta$ 가 평가항목  $ec_i$  측면에서 시점  $t$ 까지 상황  $\delta$ 에서  $o_i$ 라는 결과를 낼 확률이라고 하자. 이 확률은 기본적으로 다음과 같이 계산된다. 여기에서  $N^t(o_i)$ 는 시점  $t$ 까지 평가항목  $ec_k$  측면에서  $o_i$ 가 나타난 회수이고,  $n$ 은 전체 거래 회수이다.

$$PG^t(\alpha, \beta, \delta, o_i; ec_k) = \frac{N^t(o_i)}{n} \quad (1)$$

식(1)의 확률값은 모든 과거의 경험을 누적하여 계산하는 것으로, 개체의 행동의 개선 또는 악화가 시간의 진행에 따라 나타날 수 있음에도 불구하고, 과거 경험에 영향을 많이 받게 되는 값이다. 따라서 과거의 경험을 망각하도록 하는 것이 바람직하다. 이러한 망각효과를 확률분포에 반영하기 위해서는 시점  $t$ 에 시간구간  $[t-dt, t]$ 에 발생한 일련의 관찰결과를 순서 리스트  $OL(ec_k) = \langle (o_i, t_i) \mid t_i \in [t-dt, t] \rangle$ 에 관리한다. 여기에서  $o_i$ 는 관측값이고,  $t_i$ 는 관찰시점이다.  $PL^t(\alpha, \beta, \delta, o_i; ec_k)$ 를 시간구간  $[t-dt, t]$ 에서 평가항목  $ec_k$ 에 대한 결과  $o_i$ 의 확률이라고 하면 다음과 같이 정의할 수 있다. 여기에서  $n_{[t-dt, t]}$ 는 시간구간  $[t-dt, t]$ 에서  $\alpha$ 가  $\beta$ 와 거래한 횟수를 나타내고,  $N^{[t-dt, t]}(o_i)$ 는  $[t-dt, t]$ 내에서  $o_i$ 의 빈도수를 나타낸다.

$$PL^t(\alpha, \beta, \delta, o_i; ec_k) = \frac{N^{[t-dt, t]}(o_i)}{n_{[t-dt, t]}} \quad (2)$$

상호작용이 있을 때마다, 새로운 경험 확률값  $p^t(\alpha, \beta, \delta, o_i; ec_k)$ 은 전체 경험 확률값  $PG^t$ 과 지역 경험확률값  $PL^t(\alpha, \beta, \delta, o_i; ec_k)$ 의 가중값으로 다음과 같이 계산된다. 여기에서  $\rho$ 는 과거에 대한 망각효과를 반영하는 요소이다.

$$p^t(\alpha, \beta, \delta, o_i; ec_k) = \rho * PG^t(\alpha, \beta, \delta, o_i; ec_k) + (1 - \rho) * PL^t(\alpha, \beta, \delta, o_i; ec_k) \quad (3)$$

식(3)에 의해서 계산된 경험 확률값은 확률측도의 성질의 만족시키도록 다음과 같이 정규화시킨다.

$$P^t(\alpha, \beta, \delta, o_i; ec_k) = \frac{p^t(\alpha, \beta, \delta, o_i; ec_k)}{\sum_j p^t(\alpha, \beta, \delta, o_j; ec_k)} \quad (4)$$

### 3.1.2 만족도 계산

개체  $\alpha$ 는 각 평가항목별로 어떤 결과에 만족하는지 지정한다고 전제한다. 편의상 평가항목  $ec_i$ 에 대해서 가능한 결과가  $PO(ec_i) = \{o_{1i}, o_{2i}, \dots, o_{mi}\}$ 이다. 개체  $\alpha$ 는 각 항목별로 만족결과 집합을  $SO(\alpha, ec_i) = \{(o_1, w_1), \dots, (o_j, w_j)\}$ 와 같이 지정한다. 여기에서  $o_i \in PO(ec_i)$ 이고,  $w_{o_k} \in [0, 1]$ 은 결과  $o_k$ 에 대한  $\alpha$ 의 상대적 선호도를 나타낸다. 이때  $\alpha$ 가  $\beta$ 에 대해서  $ec_i$ 측정에 갖는 만족도는 다음과 같이 계산한다.

$$SD_\alpha(\beta, \delta; ec_i) = \sum_{(o_k, w_{o_k}) \in SO(\alpha, ec_i)} w_{o_k} \cdot P^t(\alpha, \beta, \delta, o_k; ec_i) \quad (5)$$

### 3.1.3 상황신뢰 계산

제한한 방법에서는 상황신뢰는  $\alpha$ 가  $\beta$ 에 대해서 평가항목집합  $EC$  관점에서 갖는 만족도를 이용하여 계산한다. 예를 들면, 고객이 어떤 식당에 대한 신뢰를 결정할 때, 음식의 맛, 기다리는 시간, 좋아하는 음식의 가용(availability) 여부 등 자신의 평가관점에서 신뢰를 결정한다. 제안한 방법에서  $\alpha$ 가  $\delta$ 상황에  $EC$ 관점에서  $\beta$ 에 갖는 상황신뢰는  $TS_\alpha(\beta, \delta; EC)$ 는 다음과 같이 계산한다.

$$TS_\alpha(\beta, \delta; EC) = \sum_{ec_i \in EC} w_{c_i} \cdot SD_\alpha(\beta, \delta; ec_i) \quad (6)$$

## 3.2 성향 신뢰

성향신뢰  $TD_\alpha$ 는 개체  $\alpha$ 가 다른 개체에 대해서 무조건적으로 신뢰하는 경향의 정도를 나타내는 신뢰값이다. 제안한 신뢰모델에서는 각 개체는 자신의 성향신뢰값을 지정하는 것을 전제한다. 이 성향신뢰값은 새로운 개체와 상호작용을 시작할 때, 초기 일반 신뢰값으로 사용될 수 있는 값이다.

### 3.3 일반 신뢰

일반신뢰  $TG_\alpha(\beta)$ 는 개체  $\alpha$ 가  $\beta$ 에 대해서 상황에 관계없이 갖는 신뢰의 정도이다. 일반신뢰는 새로운 상황에 대한 초기 상황신뢰 값으로 사용된다. 또한 일반신뢰는  $\beta$ 의 초기 추천(recommendation) 가중치값을 사용될 수 있다.  $\beta$ 에 대한  $\alpha$ 의 일반신뢰는 경험한 상황들  $\Phi$ 에서의 상황신뢰값을 평균하여 다음과 같이 구한다.

$$TG_\alpha(\beta) = \frac{\sum_{\delta \in \Phi} TS_\alpha(\beta, \delta; EC)}{|\Phi|} \quad (7)$$

### 3.4 명성(Reputation)

다른 개체와 거래를 처음 시작할 때, 해당 개체에 대한 명성은 매우 유용하게 사용될 수 있다. 명성은 과거의 해당 개체의 행동에 대한 관찰을 통해서, 해당 개체와 상호작용했던 개체 집단이 갖는 해당 개체에 대한 기대이다.

상황  $\delta$ 의 개체  $\beta$ 에 대한 추천자  $\gamma$ 가 있다면, 개체  $\alpha$ 는  $\beta$ 에 대한 추천정보를 이용할 수 있다. 각 개체는 자신의 평가기준에 따른 선호도에 따라 신뢰를 계산하기 때문에, 신뢰값 자체를 추천값으로 받는 것은 제안한 모델에서는 바람직하지 않다. 따라서, 제안한 모델에서는  $\beta$ 에 대한 추천자  $\gamma_j$ 의 신뢰값  $TS_{\gamma_j}(\beta, \delta; EC)$  대신 경험결과확률 분포  $P^t(\gamma_j, \beta, \delta, o_k; ec_i)$ 를 추천값으로 받는다. 추천자  $\gamma_j$ 로부터 받은 결과확률분포를 이용하여  $\gamma_j$  추천에 의한 만족도  $SD_{\gamma_j}^r(\beta, \delta; ec_i)$ 를 다음과 같이 계산한다. 여기에서는  $wo_\alpha^k$ 는 개체  $\alpha$ 가 결과  $o_k$ 에 대한 갖는 선호도이다.

$$SD_{\gamma_j}^r(\beta, \delta; ec_i) = \sum_{o_k \in SO(\alpha, ec_i)} wo_\alpha^k \cdot P^t(\gamma_j, \beta, \delta, o_k; ec_i) \quad (8)$$

추천자  $\gamma_j$ 로부터 추천에 의한 상황신뢰값은 다음

과 같이 계산한다. 여기에서  $wc_i$ 는 개체  $\alpha$ 가 평가항목  $ec_i$ 에 대해서 갖는 상대적 중요도를 나타낸다.

$$TS_{\gamma_j}^r(\beta, \delta; EC) = \sum_{ec_i \in EC} wc_i \cdot SD_{\gamma_j}^r(\beta, \delta; ec_i) \quad (9)$$

$\alpha$ 가  $\beta$ 에 대해서 갖는 추천값  $TR_\alpha(\beta, \delta; EC)$ 은 다음과 같이 추천자들  $\gamma_j$ 로 받은 상황신뢰값  $TS_{\gamma_j}^r(\beta, \delta; EC)$ 의 가중합으로 계산한다. 여기에서는 가중치  $wr_j$ 는 추천자  $\gamma_j$ 에 대해서 갖는 추천신뢰값으로,  $\alpha$ 가  $\gamma_j$ 로부터 받은 추천에 대해서 믿는 정도를 나타낸다.

$$TR_\alpha(\beta, \delta; EC) = \frac{\sum_j wr_j \cdot TS_{\gamma_j}^r(\beta, \delta; EC)}{\sum_j wr_j} \quad (10)$$

### 3.5 상황신뢰와 명성정보의 결합

개체가 어떤 커뮤니티에 참여하여 일을 하는 경우, 개체는 자신의 성향신뢰값을 설정하게 된다. 성향신뢰는 처음 어떤 개체와 거래를 시작할 때, 상대에 대한 초기 일반신뢰값으로 사용한다. 특정 상황에 대한 충분한 거래 경험이 축적되기 전까지는 일반신뢰를 상황신뢰로 사용한다. 일단 상황신뢰  $TS_\alpha(\beta, \delta; EC)$ 와 명성정보  $TR_\alpha(\beta, \delta; EC)$ 가 주어지면, 최종 신뢰값  $TS_\alpha(\beta, \delta; EC)$ 은 다음과 같이 가중합에 의해서 계산된다. 여기에서  $w \in [0, 1]$ 는 상황신뢰에 대해 상대적 중요도를 나타낸다.

$$TS_\alpha(\beta, \delta; EC) = w \cdot TS_\alpha(\beta, \delta; EC) + (1 - w) \cdot TR_\alpha(\beta, \delta; EC) \quad (11)$$

### 3.6 추천자 신뢰의 갱신

추천자  $r_i$ 의 추천신뢰도  $wr_i$ 는 추천에 의한 점수가 최종 계산된 신뢰값  $TS_\alpha(\beta, \delta; EC)$ 와 비슷한 정도에 따라 갱신된다. 추천에 의한 점수가 최종신뢰값과 비슷하다면, 추천자의 추천신뢰는 약간 증가된다. 그렇지 않으면, 추천자의 추천신뢰는 지수요소항에 의해 감소된다. 다음은 추천신뢰값 수정에 사용되는 규칙이다.  $\Delta$ 를  $|TS_\alpha(\beta, \delta; EC) - TS_{\gamma_j}^r(\beta, \delta; EC)|$ 라고 하자.  $|\Delta| < \epsilon$ 이면,  $wr_i(t+1) = \min\{wr_i(t)(1+\eta), 1\}$ 이다. 여기에서  $\epsilon$ 과  $\eta$ 는 0과 1사이의 작은 값이

다.  $\Delta$ 가  $\epsilon$ 보다 크면,  $wr_i(t+1) = wr_i(t)(1 - e^{-\gamma\Delta})$ 와 같이 갱신된다. 여기에서  $\lambda$ 는 0과 1사이의 작은 값이다.

#### 4. 결론

온라인 개체에 대한 신뢰정보는 거래를 할 때 감수해야 하는 위험을 줄일 때 매우 유용하다. 이 논문에서는 다음과 같은 특성을 갖는 계산가능한 신뢰모델을 제안하였다. 모델은 복수 개의 평가항목 측면에서 개체에 대한 신뢰도를 측정할 수 있도록 한다. 이를 위해 평가항목에 대한 경험결과분포를 축적하여, 다른 개체에 대한 신뢰도를 추정할 때, 평가개체의 경험결과에 대한 선호도를 반영할 수 있도록 하는 특징이 있다. 또한 평가항목에 대해서 다른 가중치를 부여할 수 있도록 한다. 다른 개체로부터의 추천정보를 이용할 때는, 신뢰값을 추천하도록 하는 대신, 해당 개체와의 경험결과분포 정보를 제공하도록 한다. 이를 통해 평가개체가 자신의 선호도를 추천자의 정보로부터 신뢰값을 추정하는데 반영할 수 있도록 한다.

#### 5. 참고문헌

- [1] A. Abdul-Rahman, S. Hailes. Supporting Trust in Virtual Communities. *Proc. of the Hawaii Int. Conf. on System Sciences*. (Jan.4-7, 2000, Maui Hawaii).
- [2] F. Azzedin, M. Maheswaran. Trust Modeling for Peer-to-Peer based Computing Systems. *Proc. of the Int. Parallel and Distributed Processing Symposium*. (2003).
- [3] U. Hengartner, P. Steenkiste. Implementing Access Control to People Location Information. *Proc. of SACMAT'04*. (Jun.2-4, 2004. New York).
- [4] D. Gambetta. Can We Trust Trust?. In *Trust: Making and Breaking Cooperative Relations*. Gambetta. D (ed.). Basil Blackwell. Oxford. (1990).
- [5] D. H. McKnight, N.L. Chervany. *The Meanings of Trust. Technical Report 94-04*. Carlson School of Manangement, University of Minnesota. (1996).
- [6] G. Derbas, A. Kayssi, H. artial, A. Cherhab. TRUMMAR - A Trust Model for Mobile Agent Systems Based on Reputation.

In *Proc. of ICPS2004*. IEEE. (2004).

- [7] J. Shi, G. v. Bochmann, C. Adams. A Trust Model with Statistical Foundation. In *FAST'04*. Academic Press. (2004).
- [8] Y. Wang, J. Vassileva. Bayesian Network Trust Model in Peer-to-Peer Network. In *WI'03*. IEEE. (2003).