소셜 네트워크에서 관심분야와 연결 관계를 고려한 간접 신뢰도 검출 기법

An Indirect Reliability Detection Scheme Considering Connection Relationships and Interests in Social Networks

서 인 덕*, 임 종 태**, 복 경 수**, 이 병 엽***, 유 재 수**,+
충북대학교 빅 데이터학과*, 충북대학교 정보통신공학과**,
배재대학교 사이버보안학과***

InDeok Seo*, Jongtae Lim**, Kyoungsoo Bok**,
Byoungyup Lee***, Jaesoo Yoo**
*Department of Big Data, Chungbuk National University,
**School of Information and Communication Engineering, Chungbuk National University,
***Department of Cyber Security, Paichai University

요약

소셜 네트워크 환경에서 간접 신뢰도는 직접 연결이 존재하지 않는 사용자들의 신뢰도를 평가하는데 유용한 신뢰도 평가 기법이다. 본 논문에서는 사용자간 관심분야와 연결 관계를 고려한 새로운 간접 신뢰도 검출 기법을 제안한다. 제안하는 기법은 사용자의 행위를 분석하여 관심분야를 도출한다. 또한, 멀티 홉으로 연결된 사용자간 신뢰도가 거리에 따라 변동하는 문제를 해결하기 위해 사용자간 연결 관계를 고려하여 간접 신뢰도를 검출한다.

I. 서론

인터넷과 스마트폰의 발전으로 소셜 네트워크가 활발해 지고 많은 사용자들이 활동하게 되었다. 소셜 네트워크는 사용자가 자신의 성향이나 관심분야에 대한 의견을 글로 표현 하거나 자신의 관심분야에 관련된 그룹, 페이지에 가입을 한다. 가입한 그룹이나 페이지에서 사용자간 정보 교환은 활발하게 진행된다. 그러나 이러한 활동은 사용자간 신뢰가 바탕이 되어야 활발한 상호작용이 이루어진다. 또한 신뢰성이 높은 사용자와의 교류를 해야 정확한 정보를 얻을 수 있다[1]. 사용자간 신뢰성 판별은 정확한 정보를 얻기 위한 신뢰할 수 있는 사용자를 판별하기 위해 필요하다.

신뢰도는 직접 신뢰도와 간접 신뢰도로 구별 된다. 직접 신뢰도는 사용자간 직접 연결이 되어 있는 상태에서 사용자간 상호 작용, 행위를 통해 신뢰도를 검출 한다[1, 2]. 간접 신뢰도는 사용자와 직접적으로 연결되어 있지 않은 사용자들 사이에 신뢰도이다. 사용자가 추천을 받을 때 직·간접으로 연결된 사용자에게 추천을 받는다. 신뢰도 높게 검출된 사용자에게 추천을 받는 것이 정확한 추천을 받을 수 있다. 또한 사용자간 거리가 멀어지면서 발생하는 신뢰도 손실 문제도 발생한다. 네트워크상에서 사용자간 거리가 멀티 홉 일 때 사용자 신뢰도는 사용자간 연결 관계를 고려한다[3]. 그러나 사용자의 거리로 간접 신뢰도를 검출하는 것은 한계가 있다. 사용자간 정보의 교환 시 중간 사용자의 관심분야, 성향에 의해 정확한 정보가 교환되지 않아 사용자간 정확한 신뢰도가 검출되기 어렵다.

본 논문에서는 소셜 네트워크에서 연결 관계와 관심분야를 고려한 간접 신뢰도 검출 기법을 제안한다. 제안하는 기법은 사용자의 행위를 통해 관심분야를 도출하고, 도출된 분야를 통해 사용자간 기존 신뢰도를 보정한다. 또한 사용자간 연결 관계를 고려하여 신뢰도를 검출한다. 결과적으로 제안하는 기법은 연결 관계, 관심분야 모두를 고려하기 때문에 정확한 간접신뢰도를 검출할 수 있다.

II. 제안하는 간접 신뢰도 검출 기법

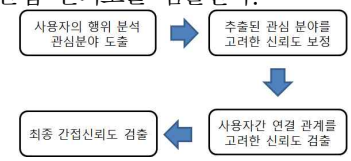
1. 제안하는 기법의 특징

본 논문은 관심분야와 사용자간의 연결 관계를 고려하여 소셜 네트워크에서 사용자간 간접 신뢰도를 검출한다. 사용자의 관심분야와 연결 관계를 고려하여 사용자간의 신뢰도를 검출 할 때 중간 사용자의 성향, 관심분야에 따라 사용자의 신뢰도가 정확하게 검출되지 않는 부분을 해결한다.

그림 1는 제안하는 전체 신뢰도 검출 과정을 나타낸다. 기반이 되는 네트워크 환경은 각 사용자간 직접 신뢰도가 계산되어 있다고 가정한다. 관심분야는 사용자가 소셜 네트워크에서 행위를 분석하여 도출한다. 도출된 관심분야는 사전에 계산되어 있던 신뢰도에 관심분야를 고려하여 사용자간 관심분야가 같을 때 사용자의 신뢰도

+ 교신저자 : yjs@chungbuk.ac.kr
이 논문은 2015년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No.2015R1D1A3A01015962), 2017년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원(No. NRF-2017R1A2B1003678), 미래창조과학부 및 정보통신기술진흥센터의 대학ICT연구센터육성 지원사업(IITP-2017-2013-0-00881)의 연구 결과로 수행되었음

를 보정 한다. 그리고 보정된 신뢰도를 기반으로 사용자의 연결 관계를 고려하여 사용자간 신뢰도를 검출한다. 사용자간 연결 홉수가 증가하는 경우 신뢰성이 저하되는 문제점을 해결한다. 앞의 두 단계에서 검출된 신뢰도를 통해 최종 간접 신뢰도를 검출한다.

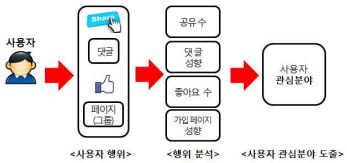


▶▶ 그림 1. 신뢰도 검출 기법

2. 관심분야 도출

소셜 네트워크에서 사용자가 자신의 성향이나 관심분야를 나태는 것은 자신이 성향이 비슷한 사용자가 있는 그룹에 가입, 친구, 제 3자의 게시물에 대한 반응, 자신의 게시물을 통해 관심분야, 성향을 알 수 있다.

본 논문에서는 그림 2와 같이 사용자의 행위를 통해 관심분야를 도출한다. 사용자의 공유 수, 댓글의 성향, 좋아요 수, 가입 페이지 성향을 분석하여 최종 사용자의 관심분야를 도출한다.

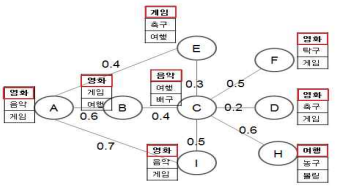


▶▶ 그림 2. 관심분야 도출

3. 관심분야를 고려한 신뢰도 보정

사용자간 관심분야가 같다면 다른 사용자에게 원하는 정보가 손실되지 않고 전달 될 수 있어 사용자간 신뢰도를 향상 시킬 수 있다.

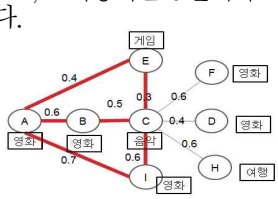
그림 3은 분야를 고려한 소셜 네트워크 환경을 나타낸다. A가 C의 간접 신뢰도를 계산하고 싶을 경우, C와 이웃한 B, E, F 등과 같은 사용자의 신뢰도 값이 활용된다. 이 단계에서는 최대 관심분야를 고려한다. A사용자 기준으로 A의 최대 관심분야는 영화이므로, 제안하는 기법은 A의 최대 관심분야인 영화를 다른 사용자의 최대 관심분야와 비교한다. 사용자간 관심분야가 같다면 C와 직접적으로 연결된 사용자 중에 관심분야가 같은 사용자의 신뢰도를 보정한다.



▶▶ 그림 3. 분야를 고려한 소셜 네트워크

4. 연결 관계를 고려한 신뢰도 검출

관심분야를 고려하여 보정 된 신뢰도를 바탕으로 연결 관계를 고려하여 신뢰도를 검출한다. 그림 4는 관계를 고려한 소셜 네트워크를 나타낸다. 이 단계는 소셜 네트워크상의 연결 관계를 고려한다. 소셜 네트워크상의 연결 관계는 사용자가 설정한 최대 홉수를 넘지 않는 연결 관계만을 고려한다. 또한 최소 홉은 2홉이고, 기준 사용자에서 타겟 사용자까지 최소 경로가 2가지 이상일 때 고려한다. 경로가 1가지 일 때는 기준 사용자와의 신뢰도를 통해 판별한다. 그림 4에서 최대 홉수가 2홉일 경우, A에서 C의 경로는 A-E-C, A-B-C, A-I-C 3가지 경로로 나타난다. 3절에서 보정된 E-C, B-C, I-C의 신뢰도와 연결 관계를 고려하여 계산한다. 각각의 경로를 통해 A에서 C까지의 신뢰도를 검출한다. A와 2홉 이상 떨어져 있는 사용자 F, D, H사용자는 3절에서 보정된 신뢰도를 그래도 유지한다.



▶▶ 그림 4. 관계를 고려한 소셜 네트워크

식(1)은 연결 관계를 고려한 신뢰도를 검출하는 식이다. 예를 들어 A-E의 신뢰도와 E-C와의 신뢰도를 통해 최종 $pt_{A_E,C}$를 검출 한다. A-E의 관계가 강한 신뢰인지 판별하기 위해 A사용자와 직접 연결된 모든 사용자의 평균을 계산한다. 평균을 계산한 후 A-E신뢰도와 차를 통해 얼마나 믿을 수 있는 사용자인지 계산한다. 그 다음 E-C의 신뢰도인 관심분야를 고려한 보정된 신뢰도와 합하여 $pt_{A_E,C}$를 검출한다. w_d는 사용자가 결정한 가중치이다. m_i는 사용자의 이웃들의 평균 신뢰도이고, $dt_{i,j}$는 i와 j의 신뢰도를 뜻한다. j는 사용자간 경로에서 중간 사용자를 나타낸다. $Npt_{j,k}$는 C의 이웃 중 j와 연결된 사용자의 분야를 고려한 보정 된 신뢰도를 뜻한다. $pt_{i_j,k}$는 연결을 고려한 신뢰도를 의미한다.

$$pt_{i_j,k} = w_d(dt_{i,j} - m_i) + Npt_{j,k} \quad (1)$$

5. 최종 간접 신뢰도 검출

앞에서 구해진 각각의 신뢰도 값을 통해 최종 신뢰도를 검출한다. 식 (3)은 각 단계에서 구해진 신뢰도를 통해 최종적인 A와 C의 최종 신뢰도 $It_{i,k}$ 값을 검출한다. $pt_{i_j,k}$는 A에서 C의 연결 관계를 j의 통해 고려한 신뢰도이고, $Npt_{l,k}$는 C의 이웃 중 A와 연결이 없는 사용자의 분야를 고려하여 보정된 신뢰도이다. 최종적으로 각 단계에서 검출된 신뢰도는 최종 계산 되어 간접 신뢰도가 검출된다.

$$It_{i,k} = \frac{\sum_{j=1}^{j} pt_{i_j,k} + \sum_{l=1}^{l} Npt_{l,k}}{n} \quad (2)$$

Ⅲ. 결론

본 논문에서는 사용자의 연결 관계와 관심분야를 고려한 간접신뢰도 검출 기법을 하였다. 향후 연구에서는 제안하는 기법의 우수성을 입증하기 위한 다양한 실험 평가를 수행할 예정이다.

■ 참 고 문 헌 ■

[1] S. Hamdi, A. L. Gançarski, A. Bouzeghoub, and S. B. Yahia, “IRIS: A Novel Method of Direct Trust Computation for Generating Trusted Social Networks,” Proc. TrustCom, pp.616-623, 2012

[2] 윤진경, 정지원, 이수지, 임종태, 복경수, 유재수, “소셜 미디어에서 사용자 행위 분석을 통한 사용자 평판 관리 기법,” 정보과학회논문지, 제43권, 제1호, pp.96-105. 2016

[3] A. Faruk and A. Paul, “A computational model for trust and reputation relationship in social network,” Proc. ICRTIT, 2016.