

Social Network 기반 Mobbing 지수 산정 알고리즘 및 분류 모델 설계

김국진, 박건우, 이상훈
국방대학교 전산정보학과
e-mail : rnrwls1977@nate.com, rnrwls1977@gmail.com

Design of Mobbing Value Computation Algorithm and Classification Model based on Social Network

Guk-Jin Kim, Gun-Woo Park, Sang-Hoon Lee
Dept of Computer Science & Information, Korea National Defense University

요 약

본 논문에서는 Mobbing(집단 따돌림) 현상에 관련된 7개의 요소(Factor)와 그 하위에 포함된 60개의 속성(Attribute)들을 선정한다. 다음으로 선정한 속성들에 대해 나와 사용자들 사이에 관계가 있으면 '1', 관계가 없으면 '0'으로 표현하고, 나와 사용자들간의 유사도 산정을 위해 각 요소안에 포함된 속성들의 합에 유사도 함수를 적용한다. 다음으로 클레멘타인의 인공신경망 알고리즘을 통해 속성들을 포함한 요소들이 취할 최적의 가중치를 산출하고, 이 값들의 총합으로 Mobbing 지수를 산정한다. 마지막으로 Social Network 사용자들의 Mobbing 지수를 본 논문에서 설계한 G2¹⁾ Mobbing 성향 분류 모델(4개의 그룹; Ideal Group of the Social Network, Bullies, Aggressive victims, Victimes)에 매핑하여 사용자들의 Mobbing 성향을 알아본다.

1. 서론

최근 청소년들을 대상으로 한 조사에 따르면 응답자의 39.5%가 따돌림을 당한 경험이 있으며, 80.4%는 학교 내 따돌림 현상이 심각하거나 매우 심각한 수준이라고 생각하는 결과가 있다[1]. 이러한 따돌림 현상의 확산은 조직 구성원의 건강에 악영향을 주고 조직과 직무에 부정적으로 작용하여 결국 사회전체에 심각한 결과를 가져온다. 이를 위해 Mobbing 현상에 대한 정확한 이해와 문제 해결을 위한 방안이 필요하다. Social Network 기반에서 Mobbing 현상은 Me-User(나와 사용자들)간에 관계유무로 표현되어 진다. 이러한 관계성을 지수화 하여 인원관리에 적용하기위해 Social Network기반 Mobbing 지수 산정 알고리즘을 제안하고 분류 모델을 설계한다. 이 알고리즘은 나와 사용자들 간의 관계를 표현하는 수개의 요소(Factor)와 속성(Attribute)들을 산정하고 관계성이 있는 경우를 '1'로 나타낸다. 이후 나와 사용자들 간의 유사도를 구하고 각 요소가 취할 최적의 가중치를 산출한다. 이렇게 개발된 알고리즘과 G2 Mobbing 성향 분류 모델을 통해 사용자를 각각의 그룹으로 분류한다.

본 논문의 구성은 2장에서는 관련연구 3장에서는 Mobbing 지수를 산정하고 4장에서는 실험 및 평가를 위한 데이터 셋과 Mobbing 지수 산정 알고리즘을 통한 분류 모델을 생성하고 마지막으로 5장에서는 결론 및 향후 연구 과제를 제시한다.

2. 관련연구

2.1 분류모델

'Michele Mouttapa'는 학생집단 내에서 Social Network 사용자를 세 가지 유형(Bullies, Aggressive victims, Victimes)으로 분류하였다[7]. 이러한 분류는 우정이나 사회계층에서 독특한 형태를 가진다. 그러므로 이 세 그룹의 행동양식이나 일반적 성향을 명확하게 정의하는 것은 중요하다. 또한 본 논문에서는 기존 연구에서 정의되지 않은 65~89%에 대한 유형을 이상적인 Social Network 사용자로 분류하고 'Ideal Group'이라 칭한다.

<표 1> Social Network 사용자의 세 가지 유형

구 분	Bullies	Aggressive victims	Victimes
학생집단의 인원비율에 따른 분류	7 ~ 15%	2 ~ 10%	2 ~ 10%

2.2 Social Network

2.2.1 Social Network 정의

Social Network은 기존의 관계를 더욱 돈독히 해주고 새로운 관계 형성을 쉽게 해주며 커뮤니케이션을 보다 빠르고 효율적으로 할 수 있도록 도와준다. 즉 사용자간의 연결이 존재한다는 것이며, 사용자는 적어도 한 가지 이상의 목적을 가지고 Social Network를 이용한다. 본격적으로 Social Network을 주목하게 만든 것은 마이스페이스와 페이스북이다. 마이스페이스(<http://kr.myspace.com>)는 전 세계 2억명이 친구로 맺어진 세상에서 가장 큰 Social Network이다. 다양한 사람들이 프로파일을 통해 취향과 관심사를 표현하고 새로운 유행과 경향을 발견하며, 정보를 공유한다.

1) G2 is the abbreviation of this Paper's two authors (Gun woo, Guk jin)

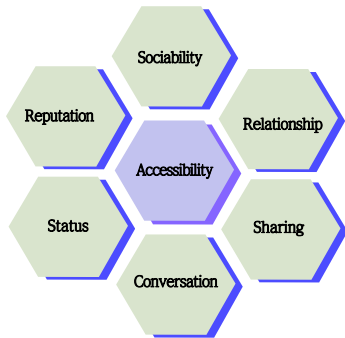
2.2.2 Social Network Service

SNS는 Social Network Service라는 뜻으로 사람들 사이의 관계, 관심에 기반한 서비스를 말한다. SNS는 블로그, UCC에 이어 웹2.0시대의 가장 주목받는 서비스로 대두되고 있다. Social Network Service의 가장 중요한 핵심은 관계(relationship)와 상호작용(interaction) 그리고 대화(communication)이다. 즉, 성공적인 SNS를 구상할 때에는 사람들 사이의 관계를 보다 쉽게 형성시켜주고, 그 관계가 상호작용하면서 관계 노드를 좀 더 넓게 만들어 주도록 해야 한다. Social Network의 유형을 분석초점에 따라 구분하면 자아중심적 네트워크, 양자 네트워크, 전체 네트워크로 구분되지만 전체 네트워크를 표현하기는 모두를 대상으로 자료를 모으기가 사실상 불가하므로 본 논문에서는 무작위로 추출된 사용자들의 자료를 통해 자아중심적 연결망(ego-centric network)을 구축한다. 즉 Me-User간 관계들에서 Me를 중심에 위치시키고 다른 사용자들과의 연결을 표현한다.

3. Mobbing 지수 산정

3.1 Mobbing 요소와 속성 산정

Erikson은 친밀성(intimacy)이란 “친밀하게 되어가는 과정에서 자신의 정체감을 잃어버리거나 앓을까 하는 두려움 없이, 타인과 개방적이고, 정신적이며, 지원적이고, 부드러운 관계를 경험할 수 있는 능력”이라고 정의하였다. 현재 지구상에 살고 있는 사람 중에서, 어떤 한 개인은 아마 그들 중의 작은 비율과 상대할 것이다. 이러한 작은 표본 중에서, 잠재적인 친구들, 적들 및 애인들은 수천명이 될 것이다. Social Network의 특징은 구성원간의 연결이 존재한다는 것이며, 이는 전술한 바와 같이 사회 현상을 투영하는 것이다. 그러므로 이러한 정의들과 사회연결망 분석의 핵심적 개념들[4], Human Network의 친밀성 요소 분석[3], Social Software Seven Building Blocks²⁾, 그 외 유수의 Mobbing 관련 논문들을 통해 Me-User간 친밀성 형성요소를 산정하기 위한 7개의 요소와 이 요소에 포함된 60개의 속성들과의 관계를 정의하였다[1]-[7].



* Peter Morville's User experience honeycomb의 변형

(그림 1) Social Software Honeycomb를 이용한 Factor관계

3.1.1 Mobbing 요소와 속성

- **Accessibility(접근성)** : 대인매력에서 나타나는 하나의 강력한

2) [KIPA '08. 7. 7; Beyond 컨슈머, Go 엔터프라이즈 Web Trend 보고서]

압력으로 지리적 근접성[3]은 유사성과 친숙성을 증가시킴 (근접성 정의요소에 물리적/기능적 거리를 포함)

- 학연, 지연(거주지), 근무연(근무지)
- 접근도(reachability ; 노드간 distance level이 1인지 2인지 즉, 관계가 직접적인지 간접적인지)
- 유사성(공통 관심사), 손쉬운 사용가능성³⁾
- 인지적 일관성(Heider의 균형이론)⁴⁾
- **Status(지위)** : 사회적 체제속에서 특정구성원이 차지하는 위치의 비교적 가치(Rayner(1997), Björkqvist et al(1994), Rayner(1997)에 의해 집단따돌림 변수로 연구)
 - 직무연관성(직무라인), 직무요구 수준
 - 직무분배/안배에 대한 부하, 직무로 인한 시간통제
- **Relationship(관계)** : 시스템 상 두 사람이 어떻게 얽혀있는지 기술하는 방식(Social Software의 Relationship 정의)
 - 내가 회원인 SNS의 회원가입, 내 블로그 방문수
 - Instant messenger 친구 등록
 - 이름, 성별, 생년
 - 인터넷 사용시 Personal Search interest
 - 취미/동호회 활동
 - 싸이월드와 같은 SNS 1촌 그룹핑
 - 스크랩, 크랙백
- **Conversation(대화)** : 시스템을 통해 다른 사람에게 대화를 건네는 방식(Social Software의 Conversation 정의)
 - 전화통화(단시간 자주, 장시간 드물게, formal, informal)
 - Instant messenger(Session기간, Conversation수)
 - e-mail 교환
- **Sociability(사회성)** : Social Network에 필요한 사회성 (Social Software의 Group 정의 및 다수의 mobbing 관련 연구의 변수)
 - 조직·자신에 대한 신념, 부하에 대한 예경
 - 적극성, 협동성, 양심성, 반응성, 융통성, 자기존중성
 - 도덕적 윤리(이기주의, 이타주의)
 - 상사와의 그룹핑[2](in-group, out-group)
 - 심리(외/내향성), 출선수범(장려, 질투/질책)
 - 판단력(냉철함, 우유부단), 사회적 관계, 의사소통, 가치명분
 - 가혹행위(신체/언어적) 받은·한 경험
 - 근무환경(괴롭힘이나 따돌림 허용/금지 분위기)
- **Reputation(평판)** : 시스템상에서 다른 사람의 배경을 통해 믿을 만한 사람이나 정보인지를 평가[5](Social software의 Reputation 정의)
 - 미니홈피/블로그 방문자수, 인터넷 커뮤니티(cafe) 참여율
 - 게시물 의 신용도(Trust list)·조회수·댓글수
 - 다수의 수신통화, 자주 전화통화 하는 사람 list
 - Instant messenger 친구 list(양방향)
 - 관리스타일(협력/고려적, 권위/방임/강압/자기과장)
 - 문제해결 방식/유형(의사존중, 독단/의존적)
- **Sharing(공유)** : 참여자들끼리 의미있는 정보를 공유하는 방식 (Social Software의 Sharing 정의)

3) 사회적 교환이론에서 알고있는 사람을 친구로 고르는 현상
4) 자신이 어울리는 사람을 좋아하려는 인지적 압력-

- 지식(재능), 보완성(나와 보완적인 사람), 코멘트
- 안부 게시판, 싸이월드 1촌평/선물/생일알림서비스
- 싸이월드 등 SNS에 방명록 작성
- SNG(Social Network Game) 참여

<표 2> 각 Factor의 Mobbing 지수 산정 테이블

구분	A ₁	A ₂	A ₃	...	A _i	Sum(A _{ij})
user1	$\sum_{j=1}^n A_{1j}$
user2	$\sum_{j=1}^n A_{2j}$
user3	$\sum_{j=1}^n A_{3j}$
user4	$\sum_{j=1}^n A_{4j}$
...
user250	$\sum_{j=1}^n A_{250j}$

3.2 Mobbing 단계 설정

Leymann(1996)은 “Bullying의 함축적 의미는 신체적 공격과 위협을 특징으로 한다. 또한 학교 따돌림에서는 자주 볼 수 있지만 직장 내 따돌림 행동에서는 신체적 공격행동이 거의 발견되지 않으며, 피해자를 사회적으로 고립시키는 등의 더욱 정교한 행동으로 특성화된다”고 하면서 이러한 성인 사회의 따돌림 행동을 Mobbing이라 칭하였다. 본 논문에서는 Mobbing을 하는 가해자의 단어가 사전에 없으므로 따돌림에 대한 가해자를 칭할 때는 Bully라는 단어를 사용한다. 학생집단 내에서 Social Network 사용자는 세가지 유형(Bullies, Aggressive victims, Victims)으로 분류된다[7]. 하지만 본 논문에서는 전술한 바와 같이 Ideal Group을 포함하여 네가지 유형으로 분류한다. Ideal Group은 조직내 65~89%가 해당하며 Social Network에 필요한 충분한 사회성을 가진 인성의 소유자이다. 다음으로 Bully는 상대방에게 반복적이고 지속적으로 신체적/감정적 위해를 가하는 자로 조직내 다른 인원들에 비해 리더십과 사회성이 높으며(Collins, Bell), 이 부류는 집단의 대략 7~15%(Pelligrini)가 해당된다. 세 번째 부류인 Aggressive victims은 Bully와 Victims의 성향을 둘 다 가진 사용자로 Bully-victims라고도 한다. 또한 감정통제가 잘 안되고, 동료간 배제나 학습에 대한 어려움을 가지며, 집단의 2~10%(Pelligrini)정도가 해당된다. 마지막 단계인 Victims은 빈번히 공격자들에 타격이 되고 거의 방어를 하지 못하는 자들로 이들은 집단의 대략 2~10%(Pelligrini)에 해당하며 예민하고 소심한 태도, 낮은 자존감으로 특징되어진다.

3.3 Mobbing 지수(MV; Mobbing Value) 산정 알고리즘

3.3.1 최적의 인성 소유자 샘플 디자인

Social Network안에 존재하는 최적의 사교성(sociability)을 가진 인성의 소유자를 만들기 위해 60개의 속성을 binary값 ‘1’로 표현하여 관계성을 부여한 후 샘플을 디자인한다. 샘플링을 하는 이유는 Social Network 사용자들과 이 샘플을 비교하여 4.2절에서 제시할 G2 Mobbing 성향 분류 모델의 지표를 산출하기 위해서이다. 최대값을 가진 샘플을 User Max (User_x)라고 한다. User_x의 Mobbing 지수는 수식(1)과 같다.

$$MV(Mobbing Value)_{User_x} = MAX(\sum_{i=1}^r A_{xi} + \sum_{j=1}^s A_{xj} + \sum_{k=1}^t A_{xk} + \sum_{l=1}^u A_{xl} + \sum_{m=1}^v A_{xm} + \sum_{n=1}^w A_{xn} + \sum_{o=1}^x A_{xu}) \quad (1)$$

여기서 $\sum_{i=1}^r A_{xi}$; 각 Factor들에 포함된 속성들의 합

3.3.2 가중치 산출

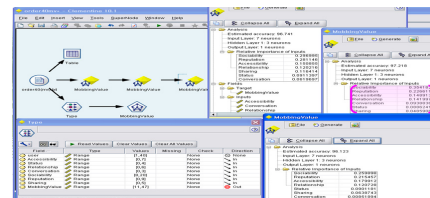
식(1)에 인공신경망(ANN; Artificial Neural Network)의 역전파 알고리즘을 적용하여 각 요소들이 가질 최적의 가중치를 산출한다. 이때 은닉노드의 활성화 함수로서 시그모이드 함수를 사용한다.

$$f(sum_j) = \frac{1}{1 + e^{-sum_j}} \quad (2)$$

이와 같이 출력층의 신경세포들도 은닉층으로부터 입력을 받아 목표 값과 활성값과의 차인 오차를 구한다. 이후 오차신호를 계산하고 신경세포 i와 j간의 연결강도들(w_{ij})을 변화시키면서 최적의 가중치를 구한다. 가중치를 식(1)에 적용하면 아래와 같다.

$$MV(Mobbing Value)_{User_x} = MAX(\alpha \sum_{i=1}^r A_{xi} + \beta \sum_{j=1}^s A_{xj} + \gamma \sum_{k=1}^t A_{xk} + \delta \sum_{l=1}^u A_{xl} + \epsilon \sum_{m=1}^v A_{xm} + \zeta \sum_{n=1}^w A_{xn} + \eta \sum_{o=1}^x A_{xu}) \quad (3)$$

(α+β+γ+δ+ε+ζ+η=1, 0 ≤ α, β, γ, δ, ε, ζ, η ≤ 1)



(그림 2) 인공 신경망을 이용한 가중치 산출

3.3.3 유사도 산출

나와 Mobbing 지수가 유사한 사용자들의 Me-User간 유사성을 확인한다. 이것은 가중치에 추가하여 Mobbing 지수(MV; Mobbing value)가 큰 경우와 특정 요소의 속성들의 합이 큰 경우가 Me-User간 유사성에 어떤 영향을 미치는지 알아보기 위함이다. 유사도 함수에는 다이스 계수(Dice's Coefficient), 자카드 계수(Jaccard's Coefficient), 코사인 계수(Cosine Coefficient) 등을 이용하는 방법이 있다. 연구 초기 코사인 계수를 이용하여 유사도를 측정하였으나, Me-User간 요소값 격차가 일정할 경우 코사인 계수가 같게 나오는 맹점이 있어 본 논문에서는 유사도 함수인 Dice 계수(Dice's Coefficient)[6]를 이용한다. Dice 계수는 문헌과 질의를 N차원 벡터값으로 표시, 문헌과 질의 벡터 간 유사도를 산출해서 검색된 문헌들을 순위화 하는 함수이다.

이 계수는 전체 요소들의 합이 큰 경우가 MV가 비슷한 사용자들 중 특정 요소의 합이 유사한 경우보다 유사도가 높게 나오는 경우를 보상에 준다.

$$SMI_{Me-User} = SMI(User_x, User_i) \quad (i = 1, 2, 3, 4, 5, \dots, n)$$

$$= \frac{2 \sum_{p=1}^y (F_{xp} \cdot F_{ip})}{\sum_{p=1}^y F_{xp}^2 + \sum_{p=1}^y F_{ip}^2} \quad (4)$$

3.3.4 Mobbing 지수 산정 알고리즘 제안

이상적인 SNS 사용자의 MV를 구함에 있어 Dice 계수를 포함시킨다면 보다 정확한 MV가 구해질 수 있다는 결론에 도달하여 Social Network 기반의 Mobbing 지수 알고리즘을 아래와 같이 제안한다.

$$MV_{Me-User} = MV(User_x, User_i)$$

$$= \frac{2 \sum_{p=1}^y (F_{xp} \cdot F_{ip})}{\sum_{p=1}^y F_{xp}^2 + \sum_{p=1}^y F_{ip}^2} \cdot User_x$$

$$MV_{Me-User} = MV(User_x, User_i)$$

$$= MAX(k \sum_{i=1}^r F_{xi}) \frac{2 \sum_{p=1}^y (F_{xp} \cdot F_{ip})}{\sum_{p=1}^y F_{xp}^2 + \sum_{p=1}^y F_{ip}^2} \cdot User_x \quad (5)$$

여기서 $MAX(k \sum_{i=1}^r F_{xi}) = User_x$

$$\sum_{p=1}^y F_{xp}^2 = Me \text{의 각 요소값들의 제곱}$$

$$\sum_{p=1}^y F_{ip}^2 = 사용자들의 각 요소값들의 제곱$$

4. 실험 및 평가

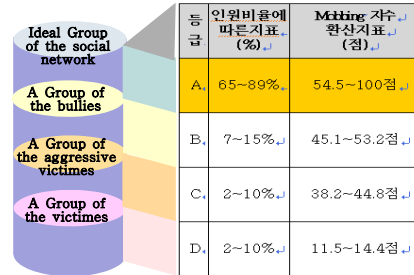
4.1 데이터 셋

제안한 알고리즘의 실험을 위해 나와 관계를 형성하고 있는 커뮤니티의 멤버들(총 40명으로 구성되며 멤버는 일반 회사원 4명, 군부대 장교 34명, 부사관 2명)을 자체 개발한 Gunwoo's Social Network Service에 가입시킨다. 실험 데이터로 사용할 40명의 사용자들과 나와 관계 정도는 추출된 속성들의 binary값으로 나타내어지며 관계가 있는 속성들('1'로 표현)의 총합으로 볼 수 있다. 멤버들 구성원 대부분이 장교 집단으로 리더십이 강하고 Social Network 형성이 원활하여, Mobbing 지수 환산지표에서 그룹별 격차(Victimes와 Aggressive victimes)가 심하게 나타남을 알 수 있다.

4.2 Mobbing 지수 산정 알고리즘을 통한 분류 모델 생성

본 논문에서는 사용자를 Mobbing 지수 산정 알고리즘의 일반화 및 실험군 형성을 위해 '학생집단'이 아닌 '일반회사(2개)', '군부대(3개 이상)'에 대해 네 가지 유형으로 분류하였다. (그림 2)는 Social Network 사용자의 G2 Mobbing 성향 분류 모델을 나타낸다. 전술한 바와 같이 기존 연구에서 분류된 세 가지 유형(Bullies, Aggressive victims,

Victimes)에 정의되지 않은 65~89%에 대한 유형을 포함하고 Ideal Group이라 칭한다. 나와 Social Network을 형성한 40명의 멤버를 제안한 알고리즘을 통해 Mobbing 지수를 산정하고 일반화한 인원비율 지표에 따라 분류하면 등급에 해당하는 Mobbing 지수 환산지표가 산출되며 이를 통해 G2 Mobbing 성향 분류 모델을 생성한다.



(그림 2) Social Network 사용자의 G2 Mobbing 성향 분류 모델

4.3 알고리즘 평가

실험집단의 Mobbing 희생자가 Mobbing 희생자로 분류될 비율을 정확률(precision)로, 알고리즘에 의해 Mobbing 희생자로 분류되어진 테스트셋 안에 있는 실제 Mobbing 희생자의 비율을 재현율(recall)로 정의하여 본 논문에서 제안한 알고리즘을 평가할 것이다.

5. 결론 및 향후 연구

Mobbing 지수는 사용자의 Mobbing 성향 파악 및 각각의 사용자들의 사회성을 확인해 볼 수 있는 지표가 된다. 제안하는 알고리즘은 현재 Mobbing 희생자뿐 만 아니라 잠재적인 Mobbing 희생자의 가능성을 파악할 수 있다. 또한 일반직장이나 군 조직의 인원관리를 위해 효과적으로 사용될 수 있다.

향후 연구계획으로는 요소와 속성을 특정집단에 적합하게 n개까지 확대시켜 나갈 것이다. 또한 제안한 알고리즘의 실험을 위해 나의 Social Network(총 500명 이상; 일반 회사원 80명, 군부대 장교 20명, 부사관 100명, 병 300명 등)을 확대해 나갈 것이다.

참고문헌

- [1] 최재승, 박경규 “조직 내 따돌림의 원인과 결과에 관한 연구-군조직을 대상으로” 2007, 서강대학교 대학원
- [2] 김광숙 “상사의 차별적 행동이 부하직원들간 상호작용에 미치는 효과” 2006, 전남대학교
- [3] 김종규, 오승환 “Human Network의 친밀성 요소 분석을 통한 정보 표현 연구” 2005, 국민대학교 대학원
- [4] 김용학 “사회 연결망 분석,” 전영사, 2003
- [5] David Kempe, Jon Kleinberg “Maximizing the Spread of Influence through a social” SIGKDD '03 Washington
- [6] Kondrak, G. “Cognates Can Improve Statistical Translation Models” in Proceedings of HLT-NAACL
- [7] Michele Mouttapa, Tom valente et al, Peggy Gallaher, Louise Ann Rohrbach, Jennifer B “Social network predictors of bullying and victimization” 2004