

온라인 소셜 네트워크에서 사용자 프로파일 기반의 모빙지수(Mobbing-Value) 알고리즘

김 국 진[†] · 박 건 우^{**} · 이 상 훈^{***}

요 약

집단 따돌림을 청소년 문제로 국한했던 것과는 달리 오늘날 직장 내 집단 따돌림은 커다란 문제로 대두되고 있다. 국제 노동기구(ILO)의 따돌림 관련 유수의 보고와 국내의 경우를 볼 때 직장 내 따돌림 경험 응답 비율이 9.1%('03)에서 30.7%('08)로 증가하고 있다. 이러한 따돌림은 개인적, 사회적으로 커다란 손실을 초래한다. 제안한 알고리즘은 사용자 프로파일을 통해 현재 Mobbing(집단 따돌림)¹⁾ 희생자뿐 만 아니라 잠재적인 Mobbing 희생자의 가능성을 파악하여 효율적인 인원관리가 가능하다.

본 논문에서는 Mobbing 현상에 관련된 사용자 프로파일 즉, 7개의 요소(Factor)와 그 하위에 포함된 50개의 속성(Attribute)들을 선정한다. 다음으로 선정한 속성들에 대해 나와 사용자들 사이에 관계가 있으면 '1', 관계가 없으면 '0'으로 표현한다. 그리고 나와 사용자들간의 유사도 산정을 위해 각 요소안에 포함된 속성들의 합에 유사도 함수를 적용한다. 다음으로 클레멘타인의 인공신경망 알고리즘을 통해 속성들이 포함된 요소가 취할 최적의 가중치를 산출하고, 이 값들의 총합으로 Mobbing 지수를 산정한다. 마지막으로 online social network 사용자들의 Mobbing 지수를 본 논문에서 설계한 G2²⁾ Mobbing 성향 분류 모델(4개의 그룹; Ideal Group of the online social network, Bullies, Aggressive victims, Victims)에 매핑하여 사용자들의 Mobbing 성향을 파악하고 이를 토대로 효율적인 인원관리에 기여할 수 있다.

키워드 : 온라인 소셜 네트워크, 모빙지수, 사용자 프로파일, G2 모빙성향 분류 모델

Mobbing-Value Algorithm based on User Profile in Online Social Network

GukJin Kim[†] · GunWoo Park^{**} · SangHoon Lee^{***}

ABSTRACT

Mobbing is not restricted to problem of young people but the bigger recent problem occurs in workspaces. According to reports of ILO and domestic case mobbing in the workplace is increasing more and more numerically from 9.1%('03) to 30.7%('08). These mobbing brings personal and social losses. The proposed algorithm makes it possible to grasp not only current mobbing victims but also potential mobbing victims through user profile and contribute to efficient personnel management.

This paper extracts user profile related to mobbing, in a way of selecting seven factors and fifty attributes that are related to this matter. Next, expressing extracting factors as '1' if they are related me or not '0'. And apply similarity function to attributes summation included in factors to calculate similarity between the users. Third, calculate optimizing weight choosing factors included attributes by applying neural network algorithm of SPSS Clementine and through this summation Mobbing-Value(MV) can be calculated. Finally by mapping MV of online social network users to G2 mobbing propensity classification model(4 Groups; Ideal Group of the online social network, Bullies, Aggressive victims, Victims) which is designed in this paper, can grasp mobbing propensity of users, which will contribute to efficient personnel management.

Keywords : Online Social Networks, Mobbing-Value, User Profile, G2 Mobbing Propensity Classification Model

[†] 준 회원 : 국방대학교 전산정보학과 석사과정
^{**} 준 회원 : 국방대학교 전산정보학과 박사과정
^{***} 종신회원 : 국방대학교 전산정보학과 교수
논문접수 : 2009년 6월 19일
수정일 : 1차 2009년 10월 28일
심사완료 : 2009년 11월 6일

1) 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 특성화 된다"고 하면서 이러한 성인 사회의 따돌림 행동을 Mobbing이라 칭하였다.

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

1. 서 론

최근 청소년들을 대상으로 한 조사에 따르면 응답자의 39.5%가 따돌림을 당한 경험이 있으며 80.4%는 학교 내 따돌림 현상이 매우 심각한 수준이라는 조사결과가 있었다. 이와 같은 따돌림 현상은 청소년 뿐만 아니라 직장내 성인들에게도 심각한 문제로 대두되고 있다. 하지만 직장내 성인들 간의 집단 따돌림에 대한 연구는 거의 보고된 것이 없는 실정이다[1]. 집단 따돌림은 집단적으로 이루어지는 동료 억압으로 건강에 심각한 영향을 줄 수 있는 사회적 스트레스의 극단적 형태이다(Olweus, 1990).

Leymann(1990)의 연구에 의하면 약 25%의 스웨덴 근로자들이, 노르웨이의 경우 전체 근로인구 중 5%가 직장내 집단 따돌림을 경험하고 있다. 영국의 경우 종업원의 53%가 작업 중에 집단 따돌림을 경험하였다고 응답하였고, 78%는 다른 직장동료가 집단 따돌림을 당하는 것을 목격했다고 했다(Rayner, 1997)[2, 7, 14, 16]. 또한 국내의 통계에서도 직장내 왕따를 당한 경험이 있다고 응답한 비율이 30.7%('08년 11월)로 보고되었다. 이러한 따돌림의 확산은 조직구성원의 건강에 악영향을 주고 조직과 직무에 부정적으로 작용하여 결국 사회전체에 심각한 결과를 가져온다는 것이 확인되고 있다. 이를 위해 모빙(Mobbing) 현상에 대한 정확한 이해와 문제 해결을 위한 방안이 필요하다. 온라인 소셜 네트워크 기반에서 Mobbing 현상은 사용자 프로파일(요소와 속성)을 토대로한 Me-User(나와 사용자들)간에 관계유무로 표현되어 진다. 이러한 관계성을 지수화 하여 인원관리에 적용하기위해 온라인 소셜 네트워크 기반 모빙 지수(Mobbing-Value) 산정 알고리즘을 제안하고 분류 모델을 설계한다. 제안하는 알고리즘은 나와 사용자들 간의 관계가 표현되는 프로파일(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. 본 논문에서 제안하는 알고리즘과 G2 Mobbing 성향 분류 모델을 통해 사용자를 각각의 그룹으로 분류하면 효율적인 인원관리가 가능하다.

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. 관련 연구

2.1 분류모델

'Michele Mouttapa'는 학생집단 내에서 Online Social Network 사용자를 세 가지 유형(Bullies, Aggressive victims, Victims)으로 분류하였다[3]. 본 논문에서는 Mobbing을 하는 가해자의 단어가 사전에 없으므로 따돌림에 대한 가해자를 칭할때는 Bully라는 단어를 사용하겠다.

약한자를 못살게 구는 공격자(Bully)는 상대방에게 반복

<표 1> Online Social Network 사용자의 네 가지 유형 분류

구 분	Ideal Group	Bullies	Aggressive victims	Victims
학생집단의 인원비율에 따른 분류	65~89%	7~15%	2~10%	2~10%

적이고 지속적으로 신체적/감정적 위해를 가하는 자이며 'Collins'와 'Bell'(1996)은 Bully들이 조직 내 다른 인원들에 비해 리더십과 사회성이 높다는 것을 발견하였다.

희생자(Victims)는 빈번히 공격자들에 타격이 되고 거의 방어를 하지 못하는 자들이며 예민하고 소심한 태도, 낮은 자존감으로 특징지어진다.

공격적 성향의 희생자(Aggressive victims)는 Bully와 victims의 성향을 둘 다가진 사용자로 Bully-Victims 라고도 한다. 이들은 공격적인 성향을 가지고 행동하기도 하나 또한 공격의 희생자가 되기도 하며 감통통제가 잘 안되고, 동료간 배제나 학습에 대한 어려움을 가진다.

본 논문에서는 이와 같은 유형 특성에 착안하여 기존에 'Michele Mouttapa'가 집단 내에서 정의하지 않은 나머지 부류(약 65~89%)를 온라인 소셜 네트워크의 이상적 그룹(Ideal Group of the Online Social Network, 이하 Ideal Group)으로 구분하여 <표 1>과 같이 Online Social Network의 사용자를 네 가지 유형으로 분류하였다.

2.2 소셜 네트워크(Social Network)

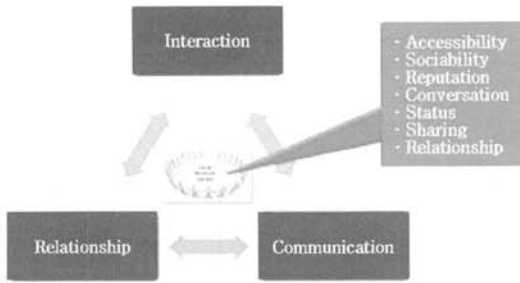
2.2.1 소셜 네트워크 정의

Online 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 Online Social Network를 이용한다.

2.2.2 소셜 네트워크 서비스

소셜 네트워크 서비스(Social Network Service, 이하 SNS)는 사람들 사이의 관계, 관심에 기반한 서비스이며 이들의 참여로 데이터가 쌓여가는 서비스를 말한다. 즉 사회적 관계개념을 인터넷 공간으로 가져온 것으로, 사람과 사람간의 관계 맺기를 통해 네트워크 형성을 지원하며 개인의 일상사나 관심사를 공유/소통 시켜주는 서비스이다. SNS는 블로그, UCC에 이어 웹2.0시대의 가장 주목받는 서비스로 대두되고 있다. 대표적인 SNS는 카페, 아이러브스쿨, 미니홈피, 마이스페이스(<http://kr.myspace.com>)와 페이스북(<http://www.facebook.com>)이다. SNS가 주목을 받는 이유는 사회속에 사는 우리가 항상 끊임없이 타인과 대화를 나누고 소통을 하려는 기본적인 욕구를 가지고 있기 때문이다.

SNS의 가장 중요한 핵심은 (그림 1)과 같이 관계와 상호작용 그리고 대화이다. 즉, 성공적인 SNS를 구성하기 위해서는 형성된 사용자간에 커뮤니케이션이 용이 하도록 해줘야한다. 그 기반하에 온라인 소셜 네트워크의 구성원들은 온라인 아이덴티티를 발전시켜가는 것이다.



(그림 1) SNS의 핵심 요소

SNS는 이미 오프라인에서 형성된 관계를 잘 관리하도록 해주는 것은 물론 (그림 2)에서와 같이 모르던 다른 사람(그룹)과의 관계를 보다 쉽게 찾아서 맺어지도록 해준다. 이것이 social network가 줄 수 있는 가치이다.

3. Mobbing 지수 산정

3.1 Mobing 요소와 속성 산정

사회 심리학자들은 핵심적인 많은 연구들에 다음의 두가지 의문을 제시한다. 첫째 “우리는 어떤 근거에서 우리의 사회적 세계를 관리 가능한 아는 이들의 숫자로 좁히게 되는가?”와 둘째 “어떤 근거에서 우리는 그들 각각을 사랑에서 미움에 이르는 매력 연속선상에 두게 되는가?” 이다. 두 사람은 그들이 단순한 물리적 근접성을 통해 함께 접촉하게 될 경우에 또한 서로간에 친숙성을 경험하고 상호간의 유사성을 가지고 있을 때 서로 좋아하고 친해진다.

Erikson은 친밀성(intimacy)이란 “친밀하게 되어가는 과정에서 자기 자신의 정체감을 잃어버리거나 앓을까 하는 두려움 없이, 타인과 개방적이고, 정신적이며, 지원적이고, 부

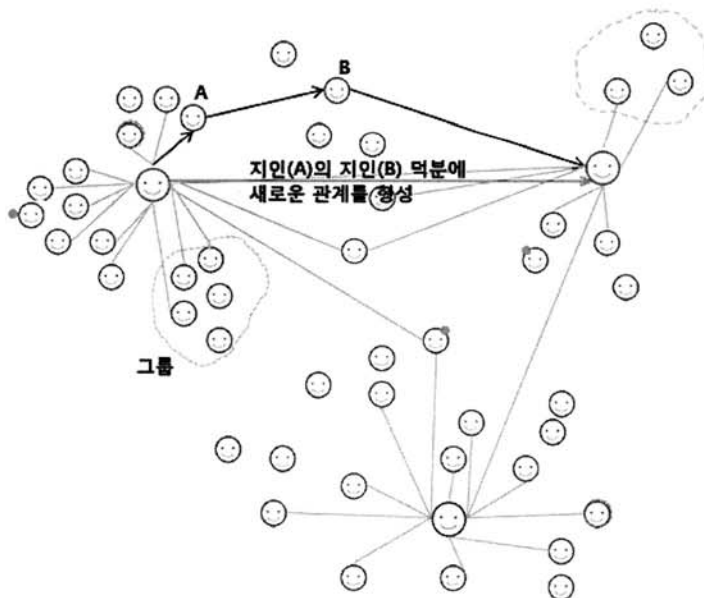
드러운 관계를 경험할 수 있는 능력”이라고 정의하였다. Stone(1973)은 인간관계에서 친밀성이란 “양자의 독립적인 판단을 지지하는 것”으로 정의하였으며, B.M. Newman과 P.R. Newman(1975)은 친밀성이란 “개인적 정체감 확립 후에 생기며, 상호감정이입(mutual empathy)과 상호욕구규정의 능력”이라고 하였다.

Online Social Network의 특징은 구성원간의 연결이 존재한다는 것이며, 이는 전술한 바와 같이 인간관계를 투영하는 것이다. 그러므로 이러한 정의들과 사회연결망 분석의 핵심적 개념[4], Human Network의 친밀성 요소 분석[5], Social Software[8] Seven Building Blocks (그림 3), 그 외 유수의 mobbing 관련 논문들을 통해 Me-User간 친밀성 형성요소를 산정하기 위한 7개의 Factor와 이 Factor에 포함된 50개의 Attribute를 <표 2>와 같이 정의하였다[11, 13-15,



* Peter Morville's User experience honeycomb의 변형

(그림 3) Social Software Honeycomb를 이용한 Factor관계



(그림 2) 모르던 사람(그룹)과의 관계형성

〈표 2〉 Mobing 요소와 속성 산정 테이블

Factor	Attribute	비고
Accessibility(접근성) : 다른 사람에서 나타나는 하나의 강력한 압력으로 지리적 근접성은 공동관심사유(사실)와 연속성을 증가시킴 [4]	나와의 확연(가끔, 보통) 나와의 지연(가끔, 보통) 나와의 은둔(가끔, 보통) 나와의 Reachable; 노드간 distance level이 '1'인지 '0'인지(직접 아는 사람, 한다리/두다리 건너 아는 사람) [5] 나와의 연지력, 열정성 : 자신과 어울리는 사람 선로 회로 성격(호감, 비호감) 나와의 유사성: 공동관심사(가끔, 보통) [5, 6] 나와의 순위를 사용가능성 : 부탁하기 쉬운(가끔, 거의 보통) 소거	* Attribute는 근접성 정의 요소에 물리적 거리적 거리를 포함
Status(지위) : 사회적 체계속에서 특정 구성원이 차지하는 위치의 비교적 가치 [7]	나와의 직무파워(유사, 공헌, 상이 경험함) 나에 대한 직무요구수준(매달, 늘어나감) 나에 대한 직무부담: 나로 인한 직무안배(매달, 직무부) 나 시간을 얼마나 통제하거나 나로 인해 시간을 통제받거나(매달, 부의함) 소거	* Attribute는 Ranker(1997), Björkqvist et al(1994) Ranker(1997)에 의해 집단 따돌림 변수로 연구
Relationship(관계) : 시스템 상 두 사람이 어떻게 연결되는지 기술하는 방식 [8]	나가가깝한 Social Network Service(SNS) 사이, 페이스북 등(가끔, 비희됨) 나의 블로그, 싸이월드, cafe 등 방문수(자주, 거의 보통) 나와의 Instant messenger 친구목록 또는 싸이월드 등 SNS에 나와의 1촌그룹(관계가끔, 관계 보통) 나와의 이름, 성명, 성명 중 한가(이성, 이성, 둘다) [9] 인터넷 사용자 나와 Personal Search interest(유사, 다름) 나와 쿼리/동호회 활동(유사, 다름) 나 블로그 쪽에 스크랩, 크린팅, 발행물 등(자주, 거의 보통) 소거	* Attribute는 Social Software의 Relationship 정의
Conversation(대화) : 시스템을 통해 다른 사람에게 대화를 건네는 방식 [9]	나와의 전화통화(단시간 자주/드물거나 자주/드물지만 사적비밀 공유, 거의 보통) 나와의 Instant messenger(문자메시지) 대화수(자주, 드물거) [10] 나와의 e-mail, 편지(e-mail 보다 대화에 높은점수 받은 편지) 교환(자주/드물지만 비밀공유, 거의 보통) 소거	* Attribute는 Social Software의 Conversation 정의
Sociability(사회성) : Online Social Network에 필요한 사회적 [11]	조직에 대한 신념(강함, 약함) 부하에 대한 예경(강함, 약함) 자신에 대한 신념(강함, 약함) 협동성(호의(미국적, 소극적) 도덕적용리(이타주의, 이기주의) 직장상사와의 그룹경(in-group/출어, out-group/관심지조) [2] 협동성(강함, 약함) 상실성(강함, 약함) 심리(희망성, 낙관성) 반응성(매달, 격명/남들) 자기존중성(강함, 약함) 운동성(가끔, 격음) 단단함(년월함, 후유부담) 가혹함(신체/언어적) 받거나 해본 경험(가끔, 격음) * 보통 '1', 격음 '0' 사회적경거(잡여겨, 고립적) 외사소통(통함, 남들) 가져업(성명, 이성됨) 은둔함(관심/취기에/취록, 격음되거나 따돌림/머물분위기) 운동성(성명, 부재/남들) 소거	* Attribute는 Social Software의 Group 정의 및 다수의 mobbing 관련 연구의 변수
Reputation(평판) : 시스템상에서 다른 사람의 태도를 알 수 있는 방식(필요한 사람이나 정보원지를 평가) [8, 11]	자신의 미니홈피/블로그를 방문하는 사람의 수 또는 Instant messenger 친구리스트 수 (가끔, 아주작음) 인터넷 커뮤니티 참여율(Cafe 등)(미국적, 소극적) 자신의 게시물 신뢰도 Trust List 또는 조회수 또는 댓글수(남들, 남들) 타인으로부터 오는 전화수(e-mail수 문자메시지 편지 수)(가끔, 거의 보통) 자신의 관리스탁(행복적/괴리적, 편위/발견/강압/자기작업) 자신의 문제해결 방식/유형 (의사판중, 독단적/의견적) 소거	* Attribute는 Social Software의 Reputation 정의
Sharing(공유) : 참여자들끼리 무엇인가 의미있는 정보를 공유하는 방식 [8]	나와의 지식/지능(공유, 미공유) 나와의 보존성/나와 보존하면 사람(강함, 거의 보통) 나와의 코멘트(공유, 미공유) [10] 나와 IQ(Social Network Game) 참여(참여, 미참여) 소거	* Attribute는 Social Software의 Shang 정의

17-20]. 이는 사용자의 프로파일을 기반으로 나와 사용자간의 관계성을 지수화하여 효율적인 인원관리에 적용하기 위함이다. 또한 (그림 3)을 통해 사용자 집단에 따른 특성화된 요소를 확인할 수 있다.

- $User_s$: 최적의 인성 소유자
- $\sum_{i=1}^r F_i$: 각 요소들에 포함된 속성들의 총합

〈표 3〉 요소와 속성 사이의 연관성 매트릭스[15]

구분	Reputation				Sociability				Accessibility				Sum(A)
	A ₁	A ₂	A ₃	A ₄	A ₁	A ₂	A ₃	A ₄	A ₁	A ₂	A ₃	A ₄	
user1													$\sum_{i=1}^r A_{1i}$
user2													$\sum_{i=1}^r A_{2i}$
user3													$\sum_{i=1}^r A_{3i}$
user4													$\sum_{i=1}^r A_{4i}$
⋮													
user1000													$\sum_{i=1}^r A_{1000i}$

* 인지적 일관성, 순위를 사용 가능성[1]

3.2 Mobing 지수(MV; Mobbing Value) 산정 알고리즘

3.2.1 최적의 인성 소유자 샘플 디자인

Online Social Network안에 존재하는 최적의 사회적(sociability)을 가진 인성의 소유자를 만들기 위해 50개의 속성을 binary 값 '1'로 표현하여 연관성을 부여한 후 샘플을 디자인한다. 샘플링을 하는 이유는 Online Social Network 사용자들과 이 샘플을 비교하여 4.2절에서 제시할 G2 Mobbing 성향 분류 모델의 지표를 산출하기 위해서이다. 최대값을 가진 샘플을 Max-User_{sample} (User_s)라고 한다. User_s의 Mobbing 지수는 식(1)과 같다.

$$\begin{aligned}
 MV(\text{Mobbing Value})_{User_s} &= MAX(\sum_{i=1}^r A_{si} + \sum_{j=1}^s A_{sj} + \sum_{k=1}^t A_{sk} + \sum_{l=1}^u A_{sl} \\
 &+ \sum_{m=1}^r A_{sm} + \sum_{n=1}^s A_{sn} + \sum_{o=1}^t A_{so}) = \sum_{i=1}^r F_i \quad (1)
 \end{aligned}$$

- 3) 사회적 교환이론에서 강조되는 상호작용의 보수들과 부담들의 균형에 영향을 준다. 즉 알고있는 사람을 친구로 고르는 현상
- 4) 우리가 싫어하는 사람의 옆에서 살거나 일하는 것은 심리적 고통이며 자신이 어울리는 사람을 좋아하려는 인지적 압력 -Heider의 균형이론

3.2.2 가중치 산출

식(1)에 인공신경망(ANN; Artificial Neural Network)의 역전과 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 오차 신호를 계산하고 신경세포 i와 j간의 연결강도들(w_{ij})을 변화시키면서 최적의 가중치를 구한다. 가중치를 식(1)에 적용하면 식(3)과 같다.

$$\begin{aligned}
 & MV(Mobbing Value)_{User_x} \\
 &= MAX(\alpha \sum_{i=1}^r A_{si} + \beta \sum_{j=1}^s A_{sj} + \gamma \sum_{k=1}^t A_{sk} + \delta \sum_{l=1}^u A_{sl} \\
 &+ \epsilon \sum_{m=1}^v A_{sm} + \zeta \sum_{n=1}^w A_{sn} + \eta \sum_{o=1}^x A_{so}) = k \sum_{i=1}^r F_i \\
 & (\alpha + \beta + \gamma + \delta + \epsilon + \zeta + \eta = 1, 0 \leq \alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta \leq 1)
 \end{aligned} \quad (3)$$

3.2.3 유사도 산출

유사도 함수인 Dice 계수(Dice's Coefficient)[12]를 이용하여 나와 Mobbing 지수가 유사한 분류 그룹 사용자들의 Me-User간 유사성을 확인한다. Me-User간 유사성을 확인하기 위한 식은 (4)와 같다.

$$\begin{aligned}
 SMI_{Me-User} &= SMI(User_x, User_i) \quad (i = 1, 2, 3, 4, 5, \dots, n) \\
 &= \frac{2 \sum_{p=1}^y (F_{xp} \cdot F_{ip})}{\sum_{p=1}^y F_{xp}^2 + \sum_{p=1}^y F_{ip}^2} \quad (4)
 \end{aligned}$$

- F_{xp} : 나의 각 요소값
- F_{ip} : 사용자들의 각 요소값

3.2.4 Mobbing 지수 산정 알고리즘 설계

Me-User간의 연관성을 기반으로 단순 Summation 식(1)에서 출발한 알고리즘에 각 요소별 가중치를 적용하면 식(3)과 같다. 또한 나와의 특정 요소에 대한 유사도 증가를 고려한 식(4)를 적용하여 알고리즘의 신뢰성을 향상시킨다. 이를 통해 보다 정확하게 Mobbing 대상자를 파악할 수 있는 Online Social Network에서 사용자 프로파일을 기반으로 한 사용자의 Mobbing 지수 산출 알고리즘을 식(5)와 같이 제안한다.

$$\begin{aligned}
 & MV(Mobbing-Value)_{User_x} \\
 &= \frac{2 \sum_{p=1}^y (F_{xp} \cdot F_{ip})}{\sum_{p=1}^y F_{xp}^2 + \sum_{p=1}^y F_{ip}^2} (k \sum_{i=1}^r F_i) \quad (5)
 \end{aligned}$$

- $k \sum_{i=1}^r F_i$: 가중치(k)가 적용된 사용자 요소값들의 총합
- $\sum_{p=1}^y F_{xp}^2$: 내 요소값들의 제곱의 합
- $\sum_{p=1}^y F_{ip}^2$: 사용자 요소값들의 제곱의 합

4. 실험 및 평가

4.1 데이터 셋

데이터셋의 표본 구성을 위해 단순임의 추출⁵⁾과 층화추출⁶⁾을 이용하였으며 단순임의 표본 추출에서는 비복원 추출에 의해 표본을 뽑는 제한적 추출방법을 선택하였다.

데이터 셋에 대한 표본 추출을 각군 병사에게 적용 시 신뢰성과 오차한계 면에서 적절한 인원을 선정하기 위해 '08년 국방백서와 '07년 9월 병력제도 개선 국무회의 자료보고를 참고하였다. 이를 토대로 구성된 각 군별 인원 비율은 아래 표와 같으나 육군의 비율이 상대적으로 커 해군과 공군은 한 부대만 표본이 되는 편의가 발생하므로 육군 60%에 해군과 공군을 각각 20%로 구성하였다.

제안한 알고리즘의 실험을 위해 나와 관계를 형성하고 있는 커뮤니티의 멤버들(총 1000명으로 구성)을 자체 개발한 Gunwoo's Social Network Service에 가입시킨다. 실험 데이터로 사용할 1,000명의 사용자들과 나와 관계 정도는 추출된 속성들의 binary값으로 나타내어지며 관계가 있는 속성들('1'로 표현)의 총합으로 볼 수 있다.

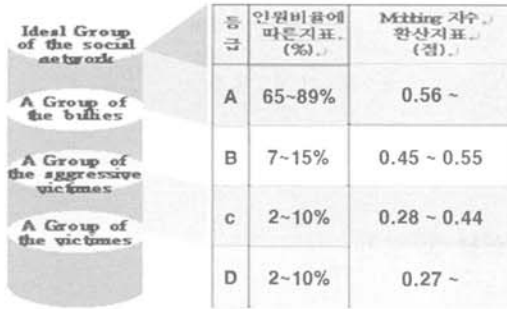
4.2 Mobbing 지수 산정 알고리즘을 통한 분류 모델 생성

본 논문에서는 사용자를 Mobbing 지수 산정알고리즘의 일반화 및 실험군 형성을 위해 '학생집단'이 아닌 '군부대'(육군 3개, 해군 2개, 공군 2개)에 적용하여 네 가지 유형으로 분류하였다. 나와 Online Social Network을 형성한 1000명의 멤버는 제안한 알고리즘을 통해 Mobbing 지수가 산정된다. 이를 일반화 한 인원비율 지표<표 4>에 따라 분류하면 등급에 해당하는 Mobbing 지수 환산지표가 산출되며 이를 통해 Online Social Network 사용자의 G2 Mobbing 성향 분류 모델을 (그림 4)와 같이 생성한다.

<표 4> 데이터 셋 구성을 위한 각 군별 인원비율

구분	장교		부사관		병		비율(%)	
	변경 전	변경 후	변경 전	변경 후	변경 전	변경 후	변경 전	변경 후
육군	80명	60명	135명	100명	561명	432명	79.6	59.2%
해군	10명	20명	18명	36명	76명	152명	10.4	20.8%
공군	10명	20명	17명	34명	73명	146명	10.0	20%
합계	100명	100명	170명	170명	730명	730명	100	100%

5) 크기가 N인 모집단에서 크기가 n인 표본을 뽑을 경우, 크기가 n인 모든 표본들이 동일한 추출 기회를 가지며 추출되는 표본추출 절차
 6) 모집단을 중복되지 않는 그룹(층)으로 나누고 각 층 으로부터 단순임의 추출로 표본을 뽑는 방법



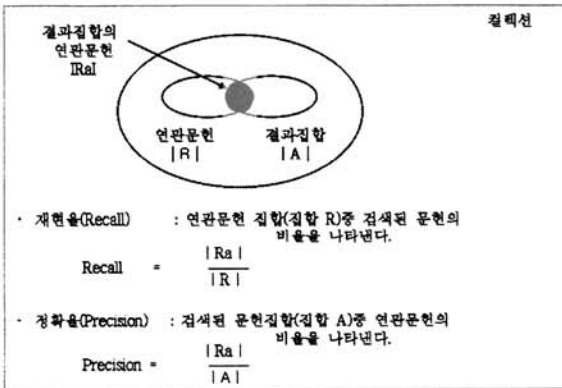
(그림 4) Online Social Network 사용자의 G2 Mobbing 상황 분류 모델

4.3 알고리즘 평가

4.3.1 재현율과 정확률

어떤 정보요구 I에 대해 연관 문헌 집합을 R이라고 가정하고, |R|은 이 집합의 문헌수를 표시한다. 어떤 검색방법이 이 정보요구를 처리하여 응답 문헌 집합 A를 검색하였다고 하고, |A|는 전과 마찬가지로 이 집합의 문헌 수를 표시한다. 또한 |R∩A|를 R과 A의 교집합의 문헌 수라고 하자. 이 집합들의 정의와 그림은 (그림 5)와 같다.

실험집단의 Mobbing 희생자가 제안하는 알고리즘에 의해 Mobbing 희생자로 분류될 비율을 정확률(precision)로, 알고리즘에 의해 Mobbing 희생자로 분류되어진 희생자가 실제 Mobbing 희생자일 비율을 재현율(recall)로 정의하여 본 논문에서 제안한 알고리즘을 평가한다.



(그림 5) 주어진 정보요구에 대한 정확률과 재현율

4.3.2 평가

알고리즘의 평가를 위해 (그림 4)의 Online Social Network 사용자 비율에 따른 지표를 평균값(mean), 최대값(Maximum, 이하 M), 최소값(minimum, 이하 m)으로 구분하고 <표 5>의 좌측과 같이 9가지 경우로 구분한다. 이렇게 구분된 9가지 경우의 비율을 데이터 셋에 적용하면 해당 인원비율에 대한 Summation값이 <표 5>의 우측과 같이 구해진다. 예를 들어 <표 5> MMM의 경우 1,000명의 데이터 셋에서 Victims 10%, Aggressive Victims 10%, Bullies 15%의 Summation을 구하면 0.311 이상, 0.312~0.503, 0.504~0.630,

<표 5> 인원비율에 따른 지표 값의 9가지 경우

구분	Victims	Aggressive victims	Bullies	Ideal Group
mmm	2	2	7	38
mMm	2	10	7	38
mmM	2	2	15	38
Mmm	2	10	15	77
MmM	10	10	15	65
MmM	10	2	7	38
MmM	10	2	15	73
MmM	10	10	7	73
mean	6	6	11	77

- M(최대값); Ideal Group의 경우 89%에 해당
- m(최소값); Ideal Group의 경우 65%에 해당
- mean(평균값); Ideal Group의 경우 77%에 해당

0.631 이상' 과 같은 지표를 얻을 수 있다.

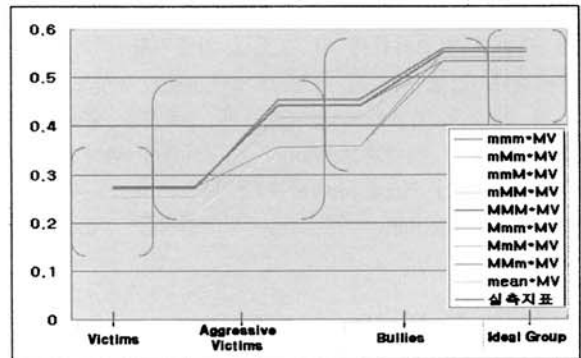
이렇게 구해진 지표에 M-V알고리즘을 적용한 후 실무부대에서 제시한 실측지표와 비교한다. <표 6(a)>는 사용자 프로파일에 대한 가중치와 유사도가 적용된 M-V알고리즘이 적용된 결과다. 이 결과 값을 그래프로 나타내면 <표 6(b)> 와 같으며 붉은색선의 실측지표와 보라색선(인원지표별 최대값에 M-V알고리즘이 적용된 경우)의 MMM*MV가 거의 유사한 경사도를 보이는 것을 확인할 수 있다.

이를 식(6) MAE(Mean Absolute Error; 절대 평균 오차)로 구해보면 단순 Summation값에 의해 구해진 지표 <표 5>의 평균은 0.045, M-V알고리즘을 적용한 지표<표 6(a)>의 평균은 0.032와 같다.

<표 6> 인원 비율에 따른 9가지 경우의 지표 비교

구분	Victimes	Aggressive victims	bullies	Ideal Group		
mmm*MV	0.21	0.22	0.36	0.37	0.53	0.54
mMm*MV	0.21	0.22	0.44	0.45	0.53	0.54
mmM*MV	0.21	0.22	0.36	0.37	0.55	0.56
Mmm*MV	0.21	0.22	0.44	0.45	0.55	0.56
MMm*MV	0.27	0.28	0.44	0.45	0.55	0.56
MmM*MV	0.27	0.28	0.36	0.37	0.53	0.54
MmM*MV	0.27	0.28	0.36	0.37	0.55	0.56
MMm*MV	0.27	0.28	0.44	0.45	0.53	0.54
mean*MV	0.25	0.26	0.42	0.43	0.54	0.55
실측지표	0.26	0.27	0.45	0.46	0.55	0.56

(a) 지표별 M-V알고리즘 적용 결과와 실무부대 실측지표 비교 결과



(b) M-V알고리즘 적용 결과의 격은 선 그래프

7) 예측 정확도를 평가하기 위한 방법으로 많은 연구에서 사용되는 방법

〈표 7〉 F-score를 통한 MMM*MV의 우수성 검증

구분	Victimes(91명)			Aggressive victims(110명)			bullies(193명)			Ideal Group(606명)		
	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score
mmm*MV	0.06593	0.06593	7%	0.01818	0.01818	2%	0.75648	0.76166	76%	0.99835	0.92409	96%
mMm*MV	0.06593	0.06593	7%	0.83636	0.87273	85%	0.75648	0.76166	76%	0.99835	0.92409	96%
Mmm*MV	0.06593	0.06593	7%	0.01818	0.01818	2%	0.87565	0.88601	88%	0.9967	0.9637	98%
mMM*MV	0.06593	0.06593	7%	0.83636	0.87273	85%	0.87565	0.88601	88%	0.9967	0.9637	98%
MMM*MV	0.84615	0.84615	85%	0.83636	0.87273	85%	0.87565	0.88601	88%	0.9967	0.9637	98%
Mmm*MV	0.84615	0.84615	85%	0.01818	0.01818	2%	0.75648	0.76166	76%	0.99835	0.92409	96%
MmM*MV	0.84615	0.84615	85%	0.01818	0.01818	2%	0.87565	0.88601	88%	0.9967	0.9637	98%
MmM*MV	0.84615	0.84615	85%	0.83636	0.87273	85%	0.75648	0.76166	76%	0.99835	0.92409	96%
mean*MV	0.63736	0.63736	64%	0.27273	0.29091	28%	0.87565	0.88601	88%	0.98185	0.9637	97%

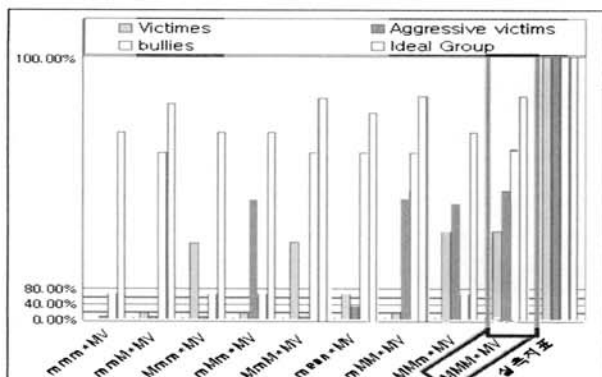
$$MAE = \frac{\sum_{i=1}^N |P_i - r_i|}{N} \quad (\text{단, } 0 \leq MAE \leq N) \quad (6)$$

- N: 예측 총 횟수
- P_i: 예측 평점
- r_i: 실제 평점

즉 MAE를 통한 예측 정확도는 MAE와 반비례하고, MAE가 낮을수록 예측 정확도가 높은 것이므로 M-V알고리즘 적용 결과가 단순 Summation보다 훨씬 더 예측 정확도가 높다는 것을 입증한다. 이점은 네 가지 분류 유형에 최대 값 비율과 M-V알고리즘을 적용한 MMM*MV가 MAE 값을 통해 다른 8가지의 경우들 보다 집단 내 각 유형별 대상자를 가장 잘 분류해 낼 수 있다는 것을 〈표 7〉을 통해 알 수 있다.

또한 F-score⁸⁾를 통해 평균값, 최대·최소값으로 구분된 9가지 경우를 비교해 보면 〈표 7〉과 같이 MMM*MV가 Online Social Network 사용자의 네 가지 유형에서 각 85%, 85%, 88%, 98%로 가장 높은 정확률을 보이는 것을 알 수 있으며, (그림 6)의 굵은 실선 부분과 같다.

그러므로 인원 비율의 최대값에 M-V알고리즘을 적용하면 〈표 8〉의 하단에서와 같이 MAE 값은 0.0067의 예측 정확률을 보이며 F-score로 확인 시에도 〈표 7〉에서와 같이 네 가지 유형특성에서 평균 89%의 일치율을 보이므로 실측 지표와 가장 유사한 지표 산출이 가능하다.



(그림 6) M-V알고리즘의 F-score

8) 정확율과 재현율을 결합한 단일 척도-최신정보검색론(김명 천 등 공역, 홍릉과학출판사)

〈표 8〉 MAE를 통한 M-V알고리즘의 예측 정확률

구분	Victimes	Aggressive victims	bullies	Ideal Group	MAE	
mmm*MV	0.21	0.22	0.36	0.37	0.53	0.05333
mMm*MV	0.21	0.22	0.36	0.37	0.55	0.04667
Mmm*MV	0.27	0.28	0.36	0.37	0.53	0.04
MmM*MV	0.27	0.28	0.36	0.37	0.55	0.03333
mMM*MV	0.21	0.22	0.44	0.45	0.53	0.02667
MMM*MV	0.21	0.22	0.44	0.45	0.55	0.02
mean*MV	0.25	0.26	0.42	0.43	0.54	0.01667
MmM*MV	0.27	0.28	0.44	0.45	0.53	0.01333
실측지표	0.26	0.27	0.45	0.46	0.55	0

5. 결론 및 향후 연구

제안한 알고리즘의 실험을 위해 나와 관계를 형성한 Community를 내가 가입한 Gunwoo's Social Network Service⁹⁾(그림 7)에 가입시켜 본 논문에서 제안한 사용자 프로파일(Factor와 Attribute)을 작성한다. 즉 모빙 관련 변수로 선정된 7개의 요소와 50개의 속성들은 Gunwoo's SNS 일반 계정에 추가되고 사용자들은 나와 연관성에 따라 관계가 있으면 '1', 관계가 없으면 '0'으로 표현되어지며 이 값들은 지수화 된다. 이를 통해 나와 연관성을 확인할 수 있다. 이것은 Mobbing Value로 지표화 될 것이며 Online Social Network 안에서 모든 사용자들의 사회성과 나와 연관성은 '0'과 '1'로 표현될 것이다.

제안하는 알고리즘은 관리자를 대상으로 개발되었으며 현재 Mobbing 희생자 뿐 아니라 잠정적인 Mobbing 희생자의 가능성을 파악할 수 있으므로 일반직장이나 군 조직의 인원 관리를 위해 효과적으로 사용될 수 있을 것이다. 또한 인성 검사¹⁰⁾와 비교해 볼 때 본 논문은 사용자들의 온라인 소셜 네트워크 성향이나 패턴이 반영되었다는데 의의가 있다고 하겠다.

향후 연구계획으로는 Factor와 Attribute를 특정집단에 맞춰 n개까지 확대시켜 나갈 것이며, Mobbing의 피해자로 분리되지 않고 희생당하는 그룹의 내재율이 최소화 될 수

9) 현재 자체 개발중인 소셜 네트워크 서비스(국방대학교, 한양 대학교, 항공대학교, 미국 리하이대학교)로 데모 버전은 거의 완성단계에 있으며, 다국적 사용자들의 참여를 위해 오프라인의 사교모임과 결합하여 가입자 수를 확대해 나가고 있음

10) 의사교환이 가능한 연령대에 그 인격을 측정하는 특성 인자로 구분된 분석법에 추가하여 정신의학적 해석방법으로 진단-길포드 박사, ERNST. KRETCHMER(독일), 김성광 교수(심리학회 이사)

있는 Factor와 Attribute를 추출할 것이다. 또한 의사교환이 가능하다고 판단하는 초등학교를 시작으로 시기와 집단에 따라 단계적이고 다계층적으로 지표를 산출하여 데이터를 표본화 한다면 보편적인 기준치를 제시할 수 있을 것이다.

참 고 문 헌

[1] 최재승, 박경규, "조직 내 따돌림의 원인과 결과에 관한 연구-군조직을 대상으로", 서강대학교대학원, pp.12-30, 2007.

[2] 박경규, "조직내 집단 따돌림과 조직행동", 제16-2집 경영논총, pp.115, 2006.

[3] Michele Mouttapa, Tom valente et al, Peggy Gallaher, Louise Ann Rohrbach, Jennifer B. "Social network predictors of bullying and victimization," University of Southern California, 39(154); pp.315-35, 2004, Summer.

[4] 김용학. "사회 연결망 분석", 전영사, 2003.

[5] 김중규, 오승환. "Human Network의 친밀성 요소 분석을 통한 정보 표현 연구", 국민대학교 대학원, pp. 10-16, 2005.

[6] HaiQiang Chen, Xueqi Cheng, Yue Liu. "Finding Core Members in Virtual Communities," pp.1233-1234, www 2008.

[7] 박경규, 최항석, "직장내 집단따돌림의 선행요인에 관한 연구", 서강대학교, 동덕여자대학교, pp.1-10, 2007.

[8] 유지은, "소셜 소프트웨어 : Beyond 컨슈머, Go 엔터프라이즈", Web Trend 보고서, KIPA, pp. 8-13, 2008년 7월호.

[9] 마크스미스 외, 조동기 역 "사이버공간과 공동체", 나남출판, pp.186, 2001.

[10] Jure Leskovec, Eric Horvitz. "Worldwide Buzz ; Planetary-Scale Views on an Instant-Messaging Network," Microsoft Research Technical Report MSR-TR-2006 -186, pp.1-5, 2007.

[11] David Kempe, Jon Kleinberg, Eva Tardos. "Maximizing the Spread of Influence through a social," SIGKDD 2003.

[12] 신동호, "Latent Semantic Analysis를 이용한 내 용기반 정보 검색 시스템", 서울대학교, pp.17-26, 2000.

[13] 김국진, 박건우, 이상훈. "Social Network 기반 Mobbing 지수 산정 알고리즘 및 분류 모델 설계", 한국 정보처리학회 춘계 세미나, 2009.

[14] 김광숙, "상사의 차별적 행동이 부하직원들간 상호작용에 미치는 효과", 전남대학교, pp.1-25, 2006.

[15] 오정운, 박건우, 이상훈. "소셜 관계 랭크 및 토픽기반-소셜 관계 랭크 알고리즘; 소셜 검색을 향해", pp.364-367, 한국정보처리학회 춘계 세미나, 2009.

[16] 최항석, 임효창, "직장내 직무따돌림이 종업원의 직무 태도에 미치는 영향", 대한경영학회 추계학술발표대회 발표논문집, pp.21-44, 2005.

[17] Dieter Zapf and Claudia Gross Johann Wolfgang Goethe-University, Frankfurt, Germany, "Conflict escalation and coping with workplace bullying; A replication and extension," EUROPEAN JOURNAL OF AND ORGANIZATIONAL PSYCHOLOGY, 10(4), pp.497-522, 2001.

[18] Dieter Zapf Johann Wolfgang Goethe-University, Frankfurt, Germany "Organisational, work group related and personal causes of mobbing/bullying at work", International Journal of Manpower, pp.70-85, 1999.

[19] Parag Singla "Yes, there is a correlation-From Social Networks to Personal Behavior on the web," pp.655-656, www2008.

[20] Zapf, Dieter; Knorz, Carmen; Kulla, Matthias " On the Relationship between Mobbing Factors, and Job Content, Social Work Environment, and Health Outcomes," European Journal of Work & Organizational Psychology. pp.215-237, 1996.



김 국 진

e-mail : rnrwls1977@nate.com

2000년 해군사관학교 전자과(학사)

2008년~현 재 국방대학교 전산정보학과 석사과정

관심분야: 정보검색, 소셜 네트워크



박 건 우

e-mail : pgw4050@hotmail.net

1997년 충남대학교 컴퓨터과학과(학사)

2007년 연세대학교 컴퓨터과학과(석사)

2008년~현 재 국방대학교 전산정보학과 박사과정

관심분야: 정보검색, 소셜 네트워크, 네트워크 보안



이 상 훈

e-mail : hoony@kndu.ac.kr

1978년 성균관대학교 정보통신공학과(학사)

1989년 연세대학교 전산학과(석사)

1997년 일본 교토대학교 정보공학 박사, 서울대학 겸임교수, 충남산업대학교 교수, 일본 교토대학교 교환교수

현 재 국방대학교 전산정보학과 교수

관심분야: 정보검색, 데이터베이스, 미디어 융합