

# 소셜 네트워크를 위한 확산 확률과 노드 연결성 기반의 정보 확산 최대화 알고리즘

응웬두이린\*, 전 문 길\*, 황 준 호\*, 유 명 식<sup>o</sup>

## An Information Diffusion Maximization Algorithm Based on Diffusion Probability and Node Degree for Social Networks

Nguyen Duy Linh\*, Wenji Quan\*, Junho Hwang\*, Myungsik Yoo<sup>o</sup>

### 요 약

최근 소셜 네트워크 서비스가 급속히 증가함에 따라 많은 기업들과 사용자들은 자신들의 정보가 더욱 빠르게 확산되길 바란다. 이러한 소셜 네트워크에서의 정보 확산을 연구하기 위하여 많은 알고리즘들이 제안되었는데 그 중 대표적인 알고리즘은 그리디 알고리즘(Greedy Algorithm)과 휴리스틱 알고리즘(Heuristic Algorithm)이다. 하지만 그리디 알고리즘의 경우 복잡성으로 인하여 실제 소셜 네트워크에 사용이 제한적이며, 기존 휴리스틱 기반의 메시지 전파 기법의 경우 균일한 소셜 네트워크 환경을 고려하였기 때문에 현재 소셜 네트워크의 특성을 반영한 연구가 요구된다. 이에 본 논문에서는 휴리스틱 알고리즘 기반의 정보 확산 확률과 노드 연결성을 고려한 정보 확산 능력 최대화 알고리즘을 제안하고, 실제 소셜 네트워크 데이터베이스를 이용하여 제안 알고리즘의 성능을 분석하였다. 그 결과 기존 알고리즘이 비해 더 많은 노드를 활성화 시킬 수 있어 정보 확산 능력을 향상시킬 수 있음을 확인하였다.

**Key Words** : Social Network, Influence, Diffusion, Heuristic Algorithm, Greedy Algorithm

### ABSTRACT

Recently, with the proliferation of social network services, users and many companies hope that their information spread more faster. In order to study the information diffusion in the social networks, many algorithms such as greedy algorithm and heuristic algorithm have been proposed. However, the greedy algorithm is too complicated to use in real-life social network, and the heuristic algorithms have been studied under the uniform distribution of diffusion probability, which is different from the real social network property. In this paper, we propose an heuristic information diffusion maximization algorithm based on diffusion probability and node degree. For performance evaluation, we use real social network database, and it is verified that our proposed algorithm activates more active nodes than existing algorithms, which enables faster and wider information diffusion.

### I. 서 론

인터넷과 웹 기술이 발전됨에 따라 최근 인터넷에서는 페이스북(Facebook), 트위터(Twitter), 카카

※ "본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음" (NIPA-2013-H0301-13-1003)

◆ 주저자 : 숭실대학교 정보통신전자공학부 통신망 설계 및 분석 연구실, linhd89@gmail.com, 정회원

◦ 교신저자 : 숭실대학교 정보통신전자공학부 통신망 설계 및 분석 연구실, myoo@ssu.ac.kr, 종신회원

\* 숭실대학교 정보통신전자공학부 통신망 설계 및 분석 연구실, moongill2008@gmail.com, 학생회원

\* 숭실대학교 정보통신전자공학부 통신망 설계 및 분석 연구실, jhwang@ssu.ac.kr, 정회원

논문번호 : KICS2013-04-188, 접수일자 : 2013년 4월 25일, 최종논문접수일자 : 2013년 5월 14일

오톡(KakaoTalk) 등과 같은 많은 소셜 네트워크 서비스(Social Network Service; SNS)들이 등장했다. 특히 스마트폰의 급격한 보급을 토대로 많은 사용자들은 소셜 네트워크 서비스를 이용하여 친구들과 소통하고 정보를 공유한다. 또한 소셜 네트워크를 이용하는 사용자들이 급속히 증가함에 따라 기업들과 광고주들의 마케팅 전략이 소셜 네트워크 흐름에 맞춰 변화되고 있다<sup>[1]</sup>.

이러한 소셜 네트워크 환경에서 현재 가장 관심을 받고 있는 것이 정보의 확산 능력(Diffusion Ability)이다. 정보의 확산 능력이란 정보의 흐름으로 표현될 수 있으며, 이 흐름의 얼마나 빨리 주변 사람들에게 확산되느냐에 따라 결정될 수 있다<sup>[2]</sup>. 즉, 사용자가 온라인상에 올린 정보(사진, 글, 동영상 등)의 빠른 시간 내에 주변 사람들에게 전파될 경우 해당 정보는 높은 확산 능력을 가졌다 할 수 있다. 일례로, 올해 4월에 발표된 싸이의 노래 '젠틀맨'의 경우 4일만에 유튜브 1억뷰를 돌파하였는데<sup>[3]</sup>, 이러한 확산 능력은 해당 정보의 영향력을 대변할 수 있는 지표라 할 수 있다.

일반적으로 소셜 네트워크 서비스는 스마트 폰을 이용한 온라인 서비스가 주를 이루기 때문에 정보의 전달을 위해서는 사용자가 이용자이면서 동시에 매개체로서의 역할을 담당한다. 따라서 정보의 확산 능력이 사용자의 확산 능력과 비례한다고 할 수 있다. 이는 해당 사용자와의 관계를 맺고 있고, 해당 정보를 볼 수 있는 사람이 많은 경우 정보가 매우 빠르게 확산될 가능성이 높기 때문이다. 하지만 사용자마다의 확산 능력은 그 사람의 직업, 나이, 사회적 위치 등의 여러 조건에 따라 각기 다르게 나타난다.

이러한 사용자들의 다른 확산 능력으로 인해 정보의 확산 능력도 다르게 형성될 수 있다. 따라서 소셜 네트워크에서 정보의 확산 능력을 향상시키기 위해서는 확산 능력이 높은 노드를 정보의 근원 노드(Seed Node)로 설정하고, 이를 정보 전달의 매개체로 삼는 것이 중요하다.

이를 위해 기존 연구에서는 그리디 알고리즘(Greed Algorithm)과 휴리스틱 알고리즘(Heuristic Algorithm)을 기반으로 소셜 네트워크 환경에서의 정보 확산을 연구하였다. 먼저, 그리디 알고리즘의 경우 확산 능력이 높은 노드를 통해 활성화 노드의 수를 증가시켜, 확산 속도를 증가시킬 수 있으나 확산 능력이 높은 노드 검출을 위해 높은 정보량과 계산량이 요구된다는 단점이 존재해 실제 소셜 네

트워크 환경에서는 사용이 제한적이다. 반면, 휴리스틱 알고리즘의 경우 그리디 알고리즘과 부분적인 정보만으로도 확산 확률을 계산할 수 있으나, 확산 속도가 그리디 알고리즘에 비해 느리다는 단점을 가지고 있다.

이에 본 논문에서는 기존 휴리스틱 알고리즘의 단점을 보완하기 위해 확산 확률(Diffusion Probability)과 각 노드의 연결성(Node Degree) 기반의 정보 확산 최대화 알고리즘을 제안한다. 또한 실제 소셜 네트워크 환경을 고려하기 위해 Netscience 데이터베이스(Database) 환경에서의 성능 평가를 수행하였다.

본 논문의 구성을 다음과 같다. 먼저 2장에서는 관련 연구로서 정보 확산의 모델 및 기존 정보 확산 기법의 장단점을 분석하고, 3장에서는 본 논문에서 제안하는 정보 확산 최대화 알고리즘을 설명한다. 이어 4장에서는 제안 알고리즘의 성능 평가 환경 및 실험 결과를 분석하고, 마지막으로 5장에서 결론을 맺는다.

## II. 관련 연구

### 2.1. 정보 확산 모델

일반적으로 소셜 네트워크 환경에서의 정보 확산 모델은 크게 선형 임계값 모델(Linear Threshold Model)과 독립 캐스케이드 모델(Independent Cascade Model)의 두 가지 모델을 기반으로 한다. 이를 보다 자세히 살펴보면 다음과 같다.

#### 2.1.1. 선형 임계값 모델

선형 임계 값 모델은 Granovetter<sup>[4]</sup>이 처음 제안하였고, Watts<sup>[5]</sup>에 의하여 일반화되었다. 이 모델은 노드들의 상태를 활성화 상태(Active State)와 비활성화 상태(Inactive State)로 구분하였다. 활성화 상태의 노드의 경우 정보의 확산 능력을 가지고 있어 주변으로부터 정보가 유입될 경우 바로 정보를 자신에게 연결된 다른 노드에게 전달하는 역할을 수행할 수 있지만, 비활성화 상태의 노드의 경우 정보 확산의 능력이 없어, 활성화 단계로의 전환(Transition)이 필요하다고 정의하고 있다.

이러한 노드의 전환 과정을 위해 선형 임계값 모델은 모든 노드의 상태를 비활성화 상태로 가정한 후 하나의 노드가 처음 정보를 채택하면서부터 상태 전환 과정을 수행한다. 즉, 하나의 노드가 활성화되면 그와 인접한 비활성화 노드에게 영향을 주

고, 이러한 영향이 축적되어 설정한 임계값을 초과하면 비활성화 노드는 활성화 상태로 전환되는 방식이라 할 수 있다.

2.1.2. 독립 캐스케이드 모델

독립 캐스케이드 모델은 확률 이론에 기반을 두어 Goldenberg<sup>[6]</sup>에 의해 제안되었다. 이 모델은 선형 임계값 모델과 같이 노드의 상태를 고려하고, 주변 노드와의 연결성을 토대로 정보 확산 능력을 분석하였다. 즉, 비활성화 노드는 인접한 활성화 노드와 일정한 확률로 연결되었는데, 이러한 확률을 비활성화 노드가 정보를 채택하는 확률로 나타내는 방법을 사용하여 정보 확산 속도를 산출하였다.

2.2. 기존 연구 분석

소셜 네트워크의 사용자들은 각기 다른 정보 확산 능력을 가지고 있다. 따라서 정보의 확산 속도 또는 확률을 증가시키기 위해 앞서 설명한 두 가지 정보 확산 모델을 기반으로 다양한 연구가 진행되었다<sup>[7-9]</sup>.

먼저, [7]에서는 확률 이론 기반의 정보의 확산 능력 분석 알고리즘을 제안하였고, [8]에서는 확산 속도 최대화 문제가 NP-hard 문제임을 증명하고, 그리디 알고리즘을 기반으로 한 최적화 솔루션을 제안하였다. 특히 [8]에서는 그리디 알고리즘의 확산 속도가 휴리스틱 알고리즘에 비해 우수하다는 것을 증명하였다. 그러나 그리디 알고리즘은 네트워크를 구성하는 모든 노드들의 확산 능력을 계산해야하고, 높은 정확도를 얻기 위해 매우 많은 반복 수행을 거치게 때문에 검출 시간 및 복잡도가 매우 높다.

이러한 그리디 알고리즘의 문제를 해결하기 위하여 [9]에서는 독립 캐스케이드 모델을 고려하여 휴리스틱 기반의 정보 확산 알고리즘(DDIC; Degree Discount IC Model)을 제안하였다. DDIC 알고리즘은 인접한 활성화 상태의 노드, 노드들 사이 연결성 그리고 각 노드가 균일 분포를 갖는 확산 확률을 고려하여 노드의 정보 확산 능력을 산출하였다. 하지만 소셜 네트워크에서 각 사용자들의 확산 확률이 모두 다르기 때문에 DDIC 기법은 실제 소셜 네트워크 환경에 적용하면 성능이 저하될 수 있다.

이와 같이 그리디 알고리즘과 휴리스틱 알고리즘은 복잡도와 정보 확산 속도 관점에서 서로 상충관계(Tradeoff)에 있다고 할 수 있다. 이에 본 논문에서는 실제 소셜 네트워크에서 가용 가능한 정보 확

산 최대화를 위해 복잡도가 낮은 휴리스틱 알고리즘을 기반으로 정보 확산 최대화 알고리즘을 제안한다.

III. 제안 정보 확산 최대화 알고리즘

3.1. 고려하는 소셜 네트워크 구조

일반적으로 소셜 네트워크 환경은 온라인에서 사용자들 관계들로 형성되었다. 이러한 관계를 그래프 구조로 전환하면 무방향 그래프로(Undirected Graph)로 볼 수 있는데, 이를  $G=(V, E, P)$ 로 표시할 수 있다. 이때  $V$ (Vertex)는 그래프의 노드의 집합,  $E$ (Edge)는 노드간 연결 링크의 집합 그리고  $P$ (Diffusion Probability)는 두 노드간의 연결된 링크의 확산 확률을 의미한다. 그림 1은 무방향 그래프 기반의 소셜 네트워크 구조의 예를 도시하고 있다. 그림에서 보는 바와 같이 노드  $v$ 를 기준으로  $V$ 는  $V=\{v, l, w, c, n\}$  와 같이 노드들의 집합으로 표시되고,  $E$ 는 노드  $v$ 와 연결된 링크의 집합  $E=\{e_{v,l}, e_{v,w}, e_{v,c}, \dots\}$  으로 표시되며,  $P$ 는 노드  $v$ 와 연결된 링크의 확산 확률의 집합  $P=\{p_{v,l}, p_{v,w}, p_{v,c}, \dots\}$  로 표시된다. 이때 확산 확률  $P$ 는 두 노드들 사이에서 정보가 성공적으로 전파되는 확률을 의미하는데, 이는 0과 1사이의 값으로 설정한다.  $P$ 의 값이 1에 근접하면 정보 확산에 성공할 확률이 높고, 0에 근접하면 성공할 확률이 낮다.

$$V = \{v, l, w, c, n\}$$

$$E = \{e_{v,l}, e_{v,w}, e_{v,c}, \dots\}$$

$$P = \{p_{v,l}, p_{v,w}, p_{v,c}, \dots\}$$

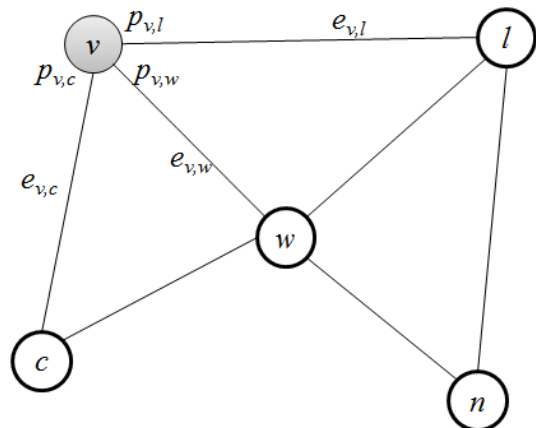


그림 1. 무방향 그래프 기반의 소셜 네트워크 구조  
Fig. 1. Structure of Social Network based on Undirected graph

이러한 소셜 네트워크 구조인 그래프 G에서의 정보 확산은 근원 노드의 집합(Seed Node Set ;  $S_{set}$ )으로부터 시작된다. 즉,  $v$ 와  $w$ 를 각각  $v \in S_{set}$ 와  $w \notin S_{set}$ 라고 설정하고, 서로 인접한 노드라고 가정한다면,  $v$ 는  $w$ 를  $P_{v,w}$ 의 확률을 가지고 활성화 시킬 수 있다. 만약 노드  $v$ 가  $w$ 의 활성화에 실패하였다면 노드  $w$ 와 인접한 다른 근원 노드가  $w$ 의 활성화를 진행한다. 이러한 노드 활성화 과정은 노드  $w$ 가 활성화 될 때까지 반복되거나  $w$ 와 연결된 모든 근원 노드들이 모두 실패할 때까지 수행된다. 이때 비활성화 노드가 활성화 노드로 전환되면, 다시 비활성화 노드로 전환되지 않는다.

### 3.2. 정보 확산 최대화 알고리즘

본 논문에서는 제안하는 정보 확산 최대화 알고리즘은 각 노드의 확산 확률과 노드 연결성을 기반으로 한다. 이를 보다 자세히 설명하면, 먼저 소셜 네트워크의 각 노드들의 확산 능력(Diffusion Ability;  $da$ )을 계산하여야한다. 이를 위해 본 논문에서는 노드 확산 능력값을 계산하기 위해 수식 1을 사용한다.

$$da_i = d_i + \alpha \sum_{l \in N(i)} p_{i,l} \quad (1)$$

이때  $da_i$ 는 노드  $i$ 의 확산 능력을 의미하고,  $d_i$ 는 노드  $i$ 의 연결성 즉 인접한 노드 수를 의미하며,  $p_{i,l}$ 는 노드  $i$ 와 노드  $l$ 간 링크의 확산 확률을 의미한다. 또한  $N(i)$ 는 노드  $i$ 와 연결된 노드의 집합을 의미하고,  $\alpha$ 는 가중치 값(Weighted Factor)을 의미한다.

이와 같은 수식에 따르면,  $da_i$ 는 노드  $i$ 와 연결된 모든 노드의 확산 확률과 연결성을 고려한 값이 된다. 이러한 각 노드의 확산 능력 산출 방법을 토대로 정보 확산 최대화를 위해서는 확산 능력이 높은 노드를 근원 노드로 결정하는 것이 바람직하다. 더욱이 소셜 네트워크에서는 한번 활성화 상태로 전환된 노드는 계속 활성화 상태로 대기해야하기 때문에 서로 인접한 근원 노드를 선택하지 않는 것이 중요하다. 이를 위해 본 논문에서는 그림 2와 같은 의사코드(Pseudo Code)를 기반으로 근원 노드 선택을 위한 알고리즘을 운용한다.

이를 상세히 살펴보면, 먼저 앞서 수식 1에서 설명한 바와 같이 네트워크를 구성하는  $n$ 개의 모든 노드들에 대한 확산 능력을 계산한다. 이후 각 노드를 확산 능력의 값에 따라 내림차순(Descending

Proposed Diffusion Maximization Algorithm	
1.	initialize $S = \emptyset$
2.	for $i=0$ to $n$ do
3.	compute its node degree $d_i$
4.	compute its proposed Diffusion Ability with $d_i$ and $p_{i,l}$ follows as eq.(1)
	$da_i = d_i + \alpha \sum_{l \in N(i)} p_{i,l}$
5.	sort array $da_i$ in descending order
6.	for $i=k$ do
7.	select $da_i = \text{argmax}\{da_i\}$
8.	if $da_i$ is not neighbor of $s$ ( $s \in S$ )
9.	$S = S \cup da_i$
10.	end for
11.	output $S$

그림 2. 제안 알고리즘의 의사코드  
Fig. 2. Pseudo Code of Proposed Algorithm

Order)으로 정렬(Sort)하고, 확산 능력이 가장 높은 노드를 근원 노드로 선택한다. 이때 한 개의 근원 노드를 선택하는 것이 아니라 해당 네트워크 구조에서 구성 가능한 최대 근원 노드를 선택하는 것이 중요하다.

이를 위해 정렬된 순서에 따라 두 노드 사이의 연결성을 비교하여 두 노드가 가까울 경우 해당 두 노드 중 확산 능력이 높은 노드를 근원 노드로 선택하고, 인접 노드는 근원 노드 선정에서 제외한다. 이러한 과정은 활성화된 모든 노드에 대한 확산 확률과 연결성 비교를 통해  $k$ 개의 근원 노드를 선택할 때 까지 반복 수행된다. 이때  $S$ 는 활성화 노드의 집합을 의미한다.

## IV. 실험 및 분석

본 논문에서는 소셜 네트워크 환경에서 정보 확산 최대화 알고리즘을 제안하였다. 이러한 제안 알고리즘의 성능 평가를 위해 1,589개의 과학자들(=노드)이 2,734개의 관계(=링크)를 형성하고 있는 실제 소셜 네트워크 데이터베이스 Netscience를 이용하였다. 이와 더불어 기존 연구와의 성능 비교를 위해 제안 알고리즘(Proposed)에서 사용하는 파라미터와 가장 근접한 특징을 갖는 기존 연구를 선택하였다. 따라서 노드의 연결성과 확산확률을 이용한 DDIC 알고리즘과 각 노드의 연결성만을 고려하는 Degree 알고리즘을 선택하고, 제안 알고리즘과의 성능을 비교 분석하였다. 표 1은 본 논문에서 고려

표 1. 실험 환경  
Table 1. Experimental Environments

Parameter	Value
Experimental computer	OS : Window 7
Program language	Visual C++ 2006
Diffusion model	Independent cascade
Diffusion distribution	Uniform, Exponential
Average diffusion probability ( $p$ )	0.1
Weighted factor ( $\alpha$ )	2

한 실험 환경을 정리한 것이다.

표에서 볼 수 있듯이 실험 환경은 윈도우 계열의 OS 환경에서 비주얼 C++를 통한 실험 환경을 구성하였다. 또한 확산 모델로 독립 캐스케이드 모델 기반의 균일 분포와 지수 분포를 가진 확산 분포를 고려하였다. 이와 더불어 각 노드의 평균 확산 확률과 조정 계수는 각각 0.1과 2로 설정하였다. 이때 균일 분포의 경우 각 노드의 전파 확률이 0.1로 고정인 반면, 지수 분포의 경우 각 노드마다 전파 확률은 다르지만 그 평균이 0.1이 되도록 설정하였다.

이와 같은 실험 환경에서 각 휴리스틱 알고리즘들의 분포 지역(Spreading Area)에서 활성화된 노드의 개수를 파악하기 위해 5,000번의 반복 실험을 수행하였고, 근원 노드  $k$ 의 수를 1 ~ 50개 까지 변화시키면서 활성화 노드의 개수를 비교 분석하였다.

그림 3은 실제 소셜 네트워크 데이터베이스인 Netscience의 노드별 전파 확률을 나타낸 것이다. 그림에서 보는 바와 같이 실제 소셜 네트워크 환경에서의 노드 분포는 지수 분포를 가지고 있다. 이는 기존 연구에서 고려한 모든 노드가 동일한 확산 확률을 가지는 균일 분포가 실제 소셜 네트워크 환경과 매우 다르다는 것을 입증하는 것이다. 더욱이 실제 소셜 네트워크 환경에서의 노드별 확산 확률 분포를 상세히 살펴보면, 0.1%의 확산 확률을 가진 노드가 전체 노드 수의 약 80% 정도를 차지하는 반면, 확산 확률 범위를 0.3% 이하에 포함되는 노드는 전체 노드의 약 98% 영역에 속하는 것을 볼 수 있다.

그림 4는 균일 분포 환경에서의 활성화 노드 수를 비교 분석한 것이다. 그림에서 보는 바와 같이 가장 기본적인 휴리스틱 모델인 Degree 알고리즘의

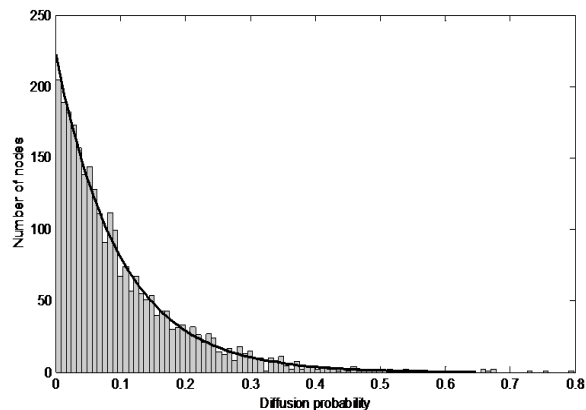


그림 3. Netscience DB에서 추출한 전파 확률에 따른 노드 분포  
Fig. 3. Node Distribution with Diffusion Probability in Netscience DB

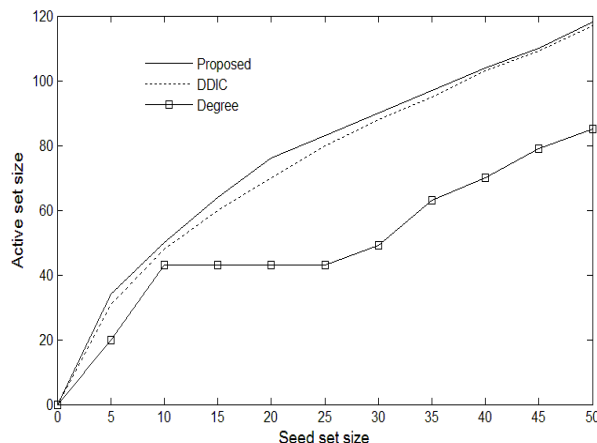


그림 4. 균일 분포하의 활성화 노드 비교  
Fig. 4. Comparison of the Number of Active Nodes under the Uniform Distribution

경우 시드 노드의 수가 증가하여도 활성화되는 노드의 수 변화가 크지 않지만, DDIC나 제안 알고리즘의 경우 시드 노드 수 증가에 비례하고 활성화 노드가 증가하는 것을 볼 수 있다.

그림 5는 지수 분포 환경에서 활성화 노드의 수를 비교한 것이다. 그림에서 보는 바와 같이 균일 분포의 성능과 유사하게 활성화 노드의 수가 증가하는 것을 볼 수 있으나, 제안 알고리즘과 DDIC 알고리즘을 비교하면, 앞서 균일 분포 하에서의 성능 차이보다 지수 분포 환경에서의 성능 차이가 크게 나타나는 것을 볼 수 있다. 특히 10개 이하의 시드 노드 환경에서는 두 알고리즘의 성능이 비슷하나, 시드 노드의 수가 증가할수록 그 성능 차이는 점차 커지며, 50개의 시드 노드가 존재할 경우 제안 알고리즘을 이용할 경우 15개 정도의 활성화 노

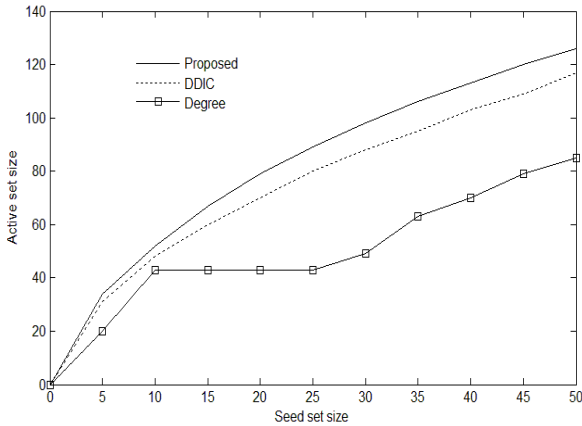


그림 5. 지수 분포하의 활성화 노드 비교  
 Fig. 5. Comparison of the Number of Active Nodes under the Exponential distribution

드가 더 검출되어 더 빠른 전파 속도를 제공할 수 있는 것으로 분석된다.

이러한 결과에서는 제안 알고리즘과 DDIC 알고리즘의 성능 차이의 원인을 상세히 분석해보면, 각 노드의 확산 확률이 동일한 균일 분포 하에서는 DDIC의 경우 확산 확률이 높은 노드만을 선택하였으나 제안 알고리즘의 경우 확산 능력이 높은 노드 중 인접한 노드를 근원 노드에서 제외시키기 때문에 보다 많은 비활성화 노드를 활성화 시킬 수 있는 것으로 분석된다. 그리고 그림 4와 5에서 볼 수 있듯이 균일 분포와 지수 분포 환경에서 제안한 알고리즘의 활성화 노드 수는 비슷한 결과로 나타나는데, 이는 균일 분포 하에서 고려한 각 노드의 평균 확산 확률이 지수 분포 하에서 80%의 노드들은 확산 확률이 비슷한 값을 갖기 때문으로 분석된다.

하지만 지수 분포 환경에서는 제안 알고리즘은 근원 노드 산정 기법과 더불어 연결된 모든 노드의 확산 확률의 합에 대한 가중치를 부여함으로써 보다 우수한 근원 노드를 산출하기 때문에 균일 분포 환경보다 더 큰 성능 차이가 발생하는 것으로 분석된다.

### V. 결 론

본 논문은 소셜 네트워크 환경에서 정보 확산 최대화에 대한 연구로서 각 노드의 확산 확률과 노드 연결성을 고려한 개선된 확산 능력 산출 방법을 제안하였다. 이와 더불어 제안 알고리즘의 성능 평가를 위해 실제 소셜 네트워크 데이터베이스인 Netscience를 이용하였고, 그 결과 기존 정보 확산

알고리즘에 비해 균일 분포와 지수 분포에서 모두 활성화 노드의 수를 증가시킬 수 있음을 확인하였다. 즉, 정보 확산 능력에 활성화 노드의 수와 비례하기 때문에 정보 확산의 최대화가 가능하다. 이러한 연구 결과는 소셜 네트워크 환경에서 복잡도는 줄이면서 정보 확산 속도를 증가시켰다는데 큰 의미를 가지고 있다고 할 수 있다.

### References

- [1] J.-W. Byeon, J.-M. Song, and D.-B. Son, "Design business model creating social network and 'self-employed' company," *J. Korean Soc. Design Sci.*, vol. 24, no. 3, pp. 195-206, Aug. 2011.
- [2] K.-S. Seol, J.-D. Kim, H.-N. Shim, and D.-K. Baik, "Intimacy measurement method and experiment between social network service users," *J. Korean Inform. Sci. Soc.*, vol. 39, no. 4, pp. 335-341, Aug. 2012.
- [3] S. Knopper, *On the Charts: Psy's 'Gentleman' Earns Youtube Love*, Retrieved Apr., 17, 2013, from <http://www.rollingstone.com>
- [4] M. Granovetter, "Threshold model of collective behavior," *The Amer. J. Sociol.*, vol. 83, no. 6, pp. 1420-1443, May 1978.
- [5] D. J. Watts, "A simple model of global cascades on random networks," in *Proc. Nat. Academy of Sci. U.S.A (PNAS)*, vol. 99, no. 9, pp. 5766-5771, April 2002.
- [6] J. Goldenberg, B. Libai, and E. Muller, "Using complex systems analysis to advance marketing theory development: modeling heterogeneity effects on new product growth through stochastic cellular automata," *Academy of Marketing Sci. Review*, vol. 2001, no. 9, pp. 1-19, 2001.
- [7] M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," in *Proc. 8<sup>th</sup> ACM Int. Conf. Knowledge Discovery and Data Mining (ACM SIGKDD)*, pp. 61-70, New York, U.S.A., August 2002.
- [8] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *Proc. 9<sup>th</sup> ACM Int. Conf.*



*Knowledge Discovery and Data Mining (ACM SIGKDD)*, pp. 137-146, New York, U.S.A., August 2003..

- [9] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in *Proc. 15<sup>th</sup> ACM Int. Conf. Knowledge Discovery and Data Mining (ACM SIGKDD)*, pp. 199-208, New York, U.S.A., June 2009.
- [10] M. E. J. Newman, "The structure and function of complex networks," *Soc. Ind. Appl. Math.*, vol. 45, no. 2, pp. 167-256, Mar. 2003.
- [11] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, "Complex networks: structure and dynamics," *Physics Reports*, vol. 424, no. 2006, pp. 175-308, Jan. 2006.

**응웬두이린 (Nguyen Duy Linh)**



2012년 6월 Hanoi University of Science and Technology 학사  
2012년 9월~현재 숭실대학교 정보통신공학과 석사과정  
<관심분야> Social Network Service, Security

**전 문 길 (Wenji Quan)**



2010년 7월 연변대학교 컴퓨터공학부 학사  
2012년 9월~현재 숭실대학교 정보통신공학과 석사과정  
<관심분야> Social Network Service, Security

**황 준 호 (Junho Hwang)**



2004년 2월 숭실대학교 정보통신전자공학부 학사  
2006년 2월 숭실대학교 정보통신전자공학부 석사  
2006년 9월~현재 숭실대학교 정보통신전자공학부 박사과정

<관심분야> Optical Access Network, Wireless MAC Protocol, Visible Light Communication, Social Network Services

**유 명 식 (Myungsik Yoo)**



1989년 2월 고려대학교 전자공학과 학사  
1991년 2월 고려대학교 전자공학과 석사  
2000년 6월 SUNY at Buffalo Dept. of EE 박사  
2000년 9월~현재 숭실대학교 정보통신전자공학부 부교수

<관심분야> Optical Network, OBS, EPON, QoS, Wireless MAC Protocol, MANET, RFID, USN, CR, Visible Light Communication, Social Network Services, Wired/Wireless Networked Control System