

다중 채널 소셜 네트워크상의 메시지 전송 모델링

(Modeling message dissemination over multi-channel social network)

김경백*

(Kyung baek Kim*)

요약

오늘날 온라인 소셜 네트워크 서비스가 대중화 되면서, 메시지 전송 연구에 있어 소셜 네트워크에 대한 역할을 이해하는 것은 중요한 문제가 되었다. 기존의 온라인 소셜 네트워크상의 메시지 전송에 대한 연구들은 주로 메시지의 확산 범위와 이를 최대화 할 수 있는 방법에 대해 다루었다. 하지만 기존의 연구들에서는 구별된 전송 특성을 가지는 다양한 채널들과 서로 다른 채널 선호도 및 재전송 특성을 가지는 소셜 네트워크 사용자의 분포가 메시지 전송에 미치는 영향에 대해서는 많이 고려하지 못했다. 이 논문에서는 이러한 다중-형태 다중-채널을 가지는 소셜 네트워크상의 메시지 전송 프로세스를 모델링하기 위해 Delay Weighted Independent Cascade 모델을 제안한다. 이 모델에서는 소셜 네트워크상의 다양한 채널들(온라인 소셜 네트워크, 이메일, SMS, 전화, 구두전달)을 고려하고 각 채널들은 서로 다른 메시지 전송 시간을 가질 수 있음을 고려하였다. 그리고, 소셜 네트워크의 각 사용자의 특성을 고려하기 위해 사용자 타입에 따라 메시지 재전송 확률 및 채널 선호도를 서로 다르게 설정하였다. 또한, 사용자의 지역 분포를 고려함으로써 다양한 상황에서의 메시지 전송 특성을 분석할 수 있도록 하였다. 제안된 모델을 기반으로 작성된 시뮬레이터를 통해, 다양한 상황의 소셜 네트워크 메시지 전송에 대해 분석하였다.

■ 중심어 : | 다중 채널 | 소셜 네트워크 | 메시지 전송 | Delay Weighted Independent Cascade

Abstract

In these days, along with the extreme popularity of online social network services, it becomes an important problem understanding the role of social network in the research of message dissemination. Past studies of message dissemination over online social network services mostly consider the coverage of message dissemination and the methods to maximize it. But, these works lack of the consideration of the impact of multi channel social network, which has multiple communication channel with distinct properties of message transfer and various users with distinct channel preferences. In this paper, the new message dissemination model over multi-modal multi-channel social network, the Delay Weighted Independent Cascade Model, is proposed. The proposed model considers various channels including online social network service, email, SMS messaging, phone and mouth-to-mouth and their distinct message transfer properties. In order to consider the various user properties, the different value of probability of forwarding a message and the different preference of communication channel is considered. Moreover, the proposed model considers the distribution of user location and allows to analyze the properties of message dissemination under various scenarios. Based on the proposed model, a message dissemination simulator is generated and the message disseminations on various scenarios are analyzed.

■ keyword : | Multi Channel | Social Network | Message Dissemination | Delay Weighted Independent Cascade

I. 서론

온라인 소셜 네트워크 (Online Social Network) 서비스는 최근 가장 각광 받는 인터넷 서비스 중 하나이다. 대표적인 온라인 소셜 네트워크 서비스로는 Facebook, Twitter, 카카오톡 등을 생각 할 수 있다. 이와 같은 온라인 소셜 네트워크의 대중

화와 함께 인터넷으로 유입되는 사용자 정보, 특히 사용자들 간의 연관 관계에 대한 정보가 크게 늘고 있다. 이 정보들을 기반으로 생성되는 소셜 네트워크 그래프는 실제 생활에서의 사람들 간의 연관 관계를 표현 할 수 있게 된다. 인터넷에 내포된 소셜 네트워크 정보는 다양한 응용 시스템에 적용된다. 소셜 네트워크 정보를 사용한 추천 시스템은 보다 효과적인 광고 전략에 사용되거나 인터넷 정보의 신뢰성을 보장하기 위해 사용될

* 정회원, 전남대학교 전자컴퓨터공학부

※이 논문은 2013년도 전남대학교 학술연구비 지원에 의하여 연구되었음

접수일자 : 2013년 06월 13일

수정일자 : 2013년 12월 12일

게재완료일 : 2013년 12월 13일

Corresponding Author : 김경백 e-mail: kyungbaekkim@jnu.ac.kr

수 있다. [1][2][4] 또한, 소셜 네트워크 정보를 사용한 메시지 전송 시스템은 보다 효과적인 뉴스 전송 시스템이나 경보 시스템에 적용 될 수 있다. [3]

소셜 네트워크 응용 시스템에서 주요 연구 이슈로는 소셜 네트워크상의 메시지의 전송 특성에 대한 이해를 생각 할 수 있다. [5][6][7] 추천 시스템의 경우 임의의 사용자가 임의의 정보에 대해서 소셜 네트워크상의 이웃 사용자들에게 추천할 경우, 얼마나 많은 사용자들에게 이 정보를 추천할 수 있는지가 주요 성능 평가지표가 된다. 즉 이전의 연구에서는 소셜 네트워크상에서 메시지 확산의 정도와 이 확산 정도를 최대화하기 위한 방법들에 초점을 맞추고 있다.

하지만, 실제 소셜 네트워크에서 메시지의 전달 과정을 살펴보면, 온라인 소셜 네트워크 서비스 뿐만 아니라 이메일, 전화, 단문 메시지 전송 등 다양한 통신 채널을 통해 메시지가 전달되는 것을 알 수 있고, 만약 사용자가 서로 가까이 위치한 경우에는 직접적으로 대화를 함으로서 메시지가 전달되는 경우도 생각 할 수 있다. 또한, 소셜 네트워크의 메시지 전송을 경보 시스템이나 공지사항 및 뉴스 전달 시스템에서 응용하려는 경우에는 단순히 메시지의 확산의 정도만을 파악하는 것 뿐 만 아니라, 다중 채널을 통해서 메시지가 전송될 때 걸리는 메시지 전달 지연 시간의 정도를 파악하는 것도 중요하게 된다.

이와 같은 고려사항들을 지원하기 위해, 이 논문에서는 다중-형태 다중-채널을 가지는 소셜 네트워크상의 메시지 전송 특성을 파악하기 위한 새로운 모델, Delay Weighted Multi Channel Independent Cascade 모델을 제안한다. 이 모델에서 소셜 네트워크 그래프의 임의의 노드는 이웃 노드 간에 다양한 타입의 통신 채널을 나타내는 다수의 에지를 가지고, 각 에지는 서로 다른 메시지 전달 특성(메시지 전달 속도, 메시지 확인 지연 시간)을 가지도록 한다. 각 사용자는 서로 다른 채널 선호도를 가질 뿐만 아니라 서로 다른 재전송 확률을 가질 수 있도록 한다. 또한 지역적 특성에 따라 특정 선호도를 가지는 사용자 분포를 조정할 수 있도록 한다. 이러한 다양한 고려사항을 토대로 구상된 Delay Weighted Multi Channel Independent Cascade 모델을 기반으로 Greedy 스타일의 메시지 전송 알고리즘을 제안 한다. 이렇게 제안된 모델을 기반으로 다중 채널 소셜 네트워크상의 메시지 전송 특성에 대한 분석을 위해 시뮬레이터를 제작하고, Facebook에서 수집한 소셜 네트워크 정보와 다양한 사용자 지역 및 유형 분포를 적용한 결과를 분석한다.

이 논문은 다음과 같이 구성되어 있다. 2장에서는 소셜 네트워크상의 메시지 전송과 그 확산 범위를 예측하기 위한 대표적 모델인 Independent Cascade 모델에 대해서 소개하고 그 한계점에 대해서 기술한다. 3장에서는 이 논문에서 제안하는 Delay Weighted Multi Channel Independent Cascade 모델과

Greedy 스타일 메시지 전송 알고리즘에 대해서 설명하고 4장에서는 메시지 전송 알고리즘에서 추가적으로 고려해야할 사항들인 사용자 유형과 지역 분포에 대해 고찰한다. 5장에서는 제안된 모델을 기반으로 구현된 시뮬레이터와 Facebook에서 추출한 소셜 네트워크와 다양한 사용자 지역 및 유형 분포를 적용한 여러 가지 메시지 전송 시나리오에 대한 분석결과를 설명한다. 마지막으로 6장에서는 이 논문의 결론을 기술한다.

II. Related Work : Independent Cascade 모델

소셜 네트워크에서 정보의 확산을 이해하기 위한 대표적인 모델은 Independent Cascade 모델이다. [7] 임의의 소셜 네트워크 그래프 $G = (V, E)$ 가 있을 때, 노드 $v \in V$ 는 소셜 네트워크 사용자를 나타내고, 에지 $e \in E$ 는 사용자간에 정보를 주고받는 채널이 존재함을 의미한다. 임의의 노드 v 는 active 상태와 inactive 상태 중 하나의 상태를 가지게 된다. 메시지 전송 시, active 상태란 노드가 메시지를 전송받았음을 의미하고, inactive 상태란 아직 노드가 메시지를 전송받지 않았음을 나타낸다. 노드는 한번 active 상태가 되면 다시 inactive 상태로 돌아오지 않고, 추가적으로 받는 메시지는 무시하게 된다.

Independent Cascade 모델에서 메시지 전송 시 inactive 상태에서 active 상태로 바뀐 노드는 받은 메시지를 소셜 네트워크상에서 통신채널이 존재하는 이웃노드들에게 전송할지에 대한 여부를 확인한다. 이를 위해 각 노드는 재전송 확률 한계 값 (P_{fwd})을 사용한다. 즉, 메시지 재전송 여부의 확인을 위해 active 상태로 바뀐 노드는 0과 1사이의 임의의 랜덤 변수의 값을 발생 시키고, 이 값이 재전송 확률 한계 값보다 작게 되면 메시지를 재전송한다. 이러한 재전송 여부에 대한 판단은 각 이웃노드들에 있어서 독립적으로 수행하게 된다. 이 재전송 과정은 active 상태로 바뀌는 노드가 더 이상 생기지 않을 때 까지 수행된다.

Independent Cascade 모델은 일반적으로 임의의 소셜 네트워크 그래프에서 정보의 확산을 최대화 할 수 있는 시드(Seed) 노드를 찾는 데 사용된다. 시드 노드란 최초로 메시지 전송을 시작하는 노드를 말한다. 정보의 확산을 최대화 하는 시드를 찾기 위한 가장 단순한 알고리즘으로는 Greedy 알고리즘을 생각할 수 있다. Greedy 알고리즘에서는 각 노드를 시드로 사용하여 Independent Cascade 모델을 통한 메시지 전송을 수행하고, 이중 메시지 전송이 끝나는 시점에서 active 상태 노드의 개수를 최대로 만드는 시드 노드를 찾는다. 이처럼 Independent Cascade 모델은 임의의 소셜 네트워크 그래프에서 메시지 전송의 확산 범위를 예측하는데 있어서 주로 사용된다. [7]

하지만, Independent Cascade 모델은 소셜 네트워크 모델을 너무 단순화함으로써, 실제 소셜 네트워크의 특성을 반영하지

못할 수 있다. 즉, 실제 소셜 네트워크에서는 사용자간에 다양한 통신 채널이 존재하지만, Independent Cascade 모델에서는 각 노드간의 에지를 하나로 설정해 사용한다. 또한, Independent Cascade 모델은 메시지 전송 시 발생하는 시간 지연에 대해 고려하지 못한다. 특히 다양한 양식을 가지는 다중 채널은 각기 그 전송 특성이 다르게 되는데, 이러한 점을 고려하기 위해서는 새로운 메시지 전송 모델이 필요하다.

III. Delay Weighted Multi Channel Independent Cascade Model

다중 채널 소셜 네트워크상에서 메시지 전송을 모델링 하기 위해 우선 다중 채널 소셜 네트워크 모델을 제시한다. 임의의 소셜 네트워크 그래프 $G = (V, E)$ 상의 임의의 두 사용자를 나타내는 두 노드 v_i, v_j 가 있을 때, 이 노드들 간에는 다수의 채널이 존재할 수 있고 이는 $e_{i,j}^{type}$ 와 같이 표현된다. 이때 가능한 type으로는 이메일(EMAIL), 온라인 소셜 네트워크 서비스(ONS), 단문 전송 메시징(SMS), 전화(PHONE) 그리고 구두 전달(MOUTH)등을 생각 할 수 있다. 이러한 두 노드간의 다중 채널 모델은 그림 1과 같다.

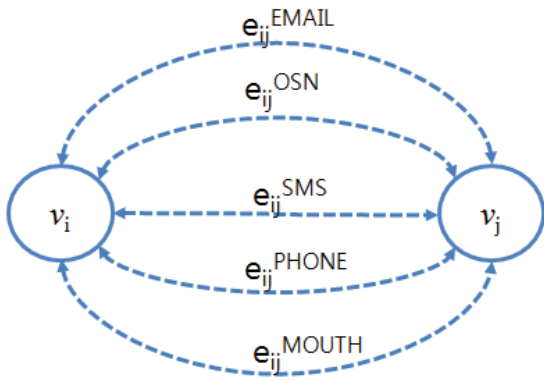


그림 1. Multi Channel Social Network Model

각 채널은 각기 서로 다른 전송 특성을 가진다. 전송 특성으로서는 메시지 전송 지연 시간(T_{delay}^{type})을 생각할 수 있는데, 이는 새로운 메시지를 통신 채널을 통해 상대방이 받을 수 있는 상태로 만드는 메시지 전달 시간과 메시지 전달이 완료된 후 메시지 수신자가 이를 확인하고 메시지의 내용을 인지하기 위해 필요한 메시지 확인 시간으로 이루어져 있다. 예를 들어, EMAIL의 경우 메시지 전달 시간은 아주 짧지만 메시지를 확인하는 시간은 상대적으로 아주 길게 된다. 반면, PHONE의 경우, 메시지 전달 시간은 상대적으로 길지만, 메시지를 확인하는 시간은 상대적으로 짧게 된다. 이와 같은 특성들을 고려해서 각기 다른 채널에서 메시지 전송 특성을 표현하게 된다.

제안되는 메시지 전송 모델에서의 메시지 재전송시, 임의의 노드 v_i 가 active상태로 바뀌게 되면 기존의 Independent Cascade 모델과 유사하게 재전송 확률 한계 값(P_{fwd})을 사용하여 재전송 여부를 확인한다. 하지만, 재전송을 해야 할 경우, 각 노드는 사용할 채널을 선택하여야 한다. 이를 위해 그림 2와 같이 각 노드는 자신의 타입에 따라 채널의 선호도를 달리 설정하고 이를 기반으로 메시지 전송 시 사용할 채널을 사용한다. 채널 선호도(P_{type})는 0보다 크고 1보다 작은 값으로 표현되고, 한 노드의 모든 채널 선호도의 합은 1이 되도록 설정한다. 즉 0부터 1사이의 임의의 균등 랜덤 값 구간을 각 채널 선호도가 대표하도록 한다. 임의의 노드는 채널 선택 시 0부터 1 사이의 랜덤 값(Uniformly Random)을 생성하고, 이 랜덤 값이 가리키는 채널 선호도 구간에 해당하는 채널을 메시지 전송 시 사용하게 된다. 예를 들어, $P_{EMAIL} = 0.2, P_{OSN} = 0.1, P_{SMS} = 0.4, P_{PHONE} = 0.3$ 과 같은 채널 선호도가 있을 경우, P_{EMAIL} 은 0부터 0.2까지의 구간, P_{OSN} 은 0.2부터 0.3까지의 구간, P_{SMS} 은 0.3부터 0.7까지의 구간, P_{PHONE} 은 0.7부터 1까지의 구간의 균등 랜덤 값 구간을 대표하도록 하고, 만약 임의의 노드가 0.55의 랜덤 값을 생성 하였을 경우, SMS를 메시지 전송 채널로 사용하도록 한다.

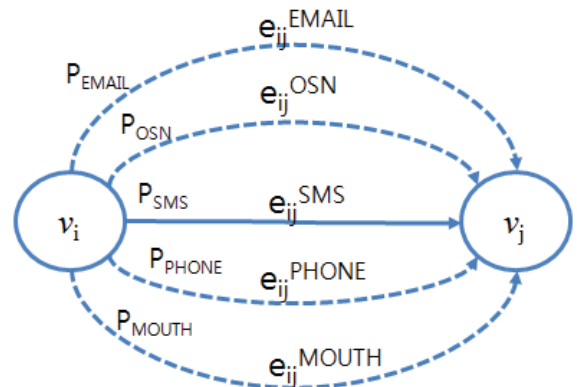


그림 2. Preferred Channel Selection

이러한 채널 중 구두로 전송되는 MOUTH채널의 경우는 다른 채널들과는 달리 사용자들 간의 장소정보에 영향을 많이 받게 된다. 즉 사용자들이 서로 가까운 곳에 위치하고 있을 경우 메시지 전송은 사람들 간의 대화에 의해서 즉각적으로 이루어지게 되고 특정 경우에는 재전송 여부 및 채널 선택 여부 등을 수행할 필요가 없게 된다. 이를 위해 제안된 메시지 전송 모델에서는 P_{MOUTH} 값이 특별히 정해져 있지 않는 경우에는 사용자 간의 거리가 일정 한계 값($Thres_{dist}$)보다 작을 경우, 재전송 확률(P_{fwd})을 1로 하고, 전송 지연 시간(T_{delay}^{type})을 아주 작은 값으로 설정한다. 반면, P_{MOUTH} 이 정해져 있을 경우, MOUTH

채널이 전송 채널로 선택되는 경우에는 전송 지연 시간(T_{delay}^{type}) 이 사용자간의 거리에 따라 선형적으로 비례해 증가하도록 설정하였다.

이러한 다중채널 소셜 네트워크상의 다양한 타입의 사용자를 고려한 메시지 전송 모델을 통한 메시지 정보 확산 정도를 확인하기 위한 Greedy 스타일 알고리즘은 그림 3과 같다. 이 알고리즘을 통해 다중 채널 소셜 네트워크상에서 메시지 전송의 확산정도 뿐만 아니라 메시지 전송 시 발생하는 메시지 전송 지연 시간의 정도에 대해서도 확인 할 수 있다.

```

QUEUE → time sorted queue
v_s → random seed

v_s.setMessageTime(0)
Queue.add(v_s)
while( Queue.hasNext() ) {
    v_i = Queue.pop()
    t_cur = v_i.getMessageTime()
    for each v_j in v_i.neighbors {
        if( v_j.isInactive() ) {
            if( v_i.P_MOUTH == 0
                && distance(v_i, v_j) < Thres_dist ) {
                v_j.setMessageTime( t_cur + T_MOUTH )
            }
            else {
                if( v_i.P_fwd > random() ) {
                    type = v_i.selectPreferredChannel()
                    delay = v_i.getDelayOfType(type)
                    v_j.setMessageTime( t_cur + delay )
                    Queue.add(v_j)
                }
            }
        }
    }
}

```

그림 3. Pseudo code of greedy algorithm of message dissemination with delay weighted multi channel independent cascade model

IV. 사용자 유형과 지역 분포

제안된 Delay Weighted Multi Channel Independent Cascade 모델을 사용하여 메시지 전송을 시뮬레이션하기 위해서는 사용자의 유형과 사용자의 지역 분포를 고려하여야 한다. 앞장에서 언급한 것과 같이 각 사용자들은 자신들만의 통신

채널 선호도를 가지게 된다. 예를 들어 청소년들은 카카오톡, Facebook 등 온라인 소셜 네트워크를 주로 사용하여 메시지를 주고받는 반면 컴퓨터 시스템에 익숙하지 못한 시니어 분들은 전화나 SMS를 주로 사용하게 된다. 이러한 점을 고려하여, 현재 구현된 시뮬레이터에서는 침대, 컴퓨터 전공 대학생, 가정주부, 시니어, 직장인의 총 다섯 가지의 사용자 유형을 설정했다. 또한, 임의의 사용자는 속해 있는 사용자 유형에 따라 서로 다른 채널 선호도뿐만 아니라 서로 다른 재전송 확률을 가지게 된다. 스포츠에 관련된 메시지에는 침대나 직장인들이 높은 관심을 보이며 높은 메시지 재전송 확률을 가지는 것과는 달리 생활 및 건강정보에 관련된 메시지에는 가정주부나 시니어들이 높은 관심을 보이게 된다.

또한, 소셜 네트워크 사용자의 분포는 지리적 특성과 관련을 가질 수 있다. 일과 시간에는 대부분의 사람들이 직장에서 일하게 되어 도심에 사용자들이 많이 분포하는 반면, 저녁 시간에는 사용자들이 거주 지역에 분포하게 되어 상대적으로 넓게 퍼져서 사용자들이 위치하게 된다. 다양한 지역 분포를 표현하기 위해서, 그림 4와 같은 sparse, urban, two cities, three cities 와 같은 다양한 형태의 사용자 지역 분포를 고려한다. Sparse의 경우는 사용자들이 균등하게 전 지역에 퍼져 있는 경우를 나타낸다. Urban의 경우는 특정지역에 많은 사용자들이 몰리는 것을 표현한 것으로, 이를 위해 특정 지역의 중심 좌표를 기준으로 가우시안 (Gaussian) 분포로 사용자들을 위치시킴으로써 구체화 한다. Two cities나 Three cities의 경우는 사용자들이 여러 지역에 몰리는 경우를 표현하기 위해 사용된다. 이러한 Urban, Two cities, Three cities을 편집된 (skewed) 분포라고 말할 수 있고, 이 밀집 지역은 각각 지역적 특색을 가질 수 있다. 예를 들어 위치하는 지역이 금융권이 밀집해 있는 도심이라면 이 지역에 위치하는 사용자들은 대부분이 직장인이 될 것이다.

이러한 두 가지 분포(사용자 유형 분포와 사용자 지역 분포)는 소셜 네트워크 사용자들의 분포를 고려할 때 동시에 고려되어야 한다. 즉 사용자 지역분포가 고려된 후 사용자의 유형분포를 고려해주어야 한다는 것이다. 예를 들어 사용자들의 지역분포는 sparse할 때, 이 지역의 사용자 유형분포로 균등 분포를 사

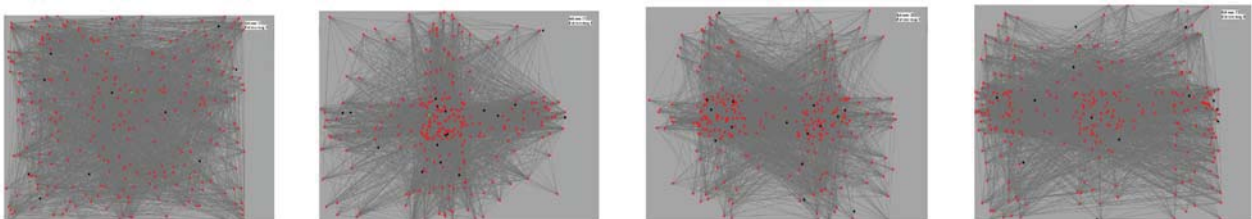


그림 4. 사용자 지역 분포. 각 점은 사용자를 나타내고, 선들은 통신 채널을 나타낸다. 왼쪽부터:sparse분포, urban분포, two cities 분포, three cities 분포

용할 경우 모든 유형의 사용자들이 균등하게 퍼지게 되고, 이 지역의 사용자 유형분포를 직장인을 위주로 분포시킨다면 대부분의 사용자들이 직장인 유형을 가지게 된다. 이와 같이 두 개의 사용자 분포를 동시에 고려함으로써 제안된 모델이 다양한 형태의 소셜 네트워크상의 메시지 전송을 분석할 수 있게 한다.

V. 다중 채널 소셜 네트워크 메시지 전송 분석

멀티 채널 소셜 네트워크 상에서의 메시지 전송 특성 분석을 위해 제안된 Delay Weighted Multi Channel Independent Cascade 모델과 Greedy 스타일 메시지 전송 알고리즘을 기반으로 시뮬레이터를 제작하였다. 구현된 시뮬레이터는 이벤트 기반의 시뮬레이터로 매 틱(tick)마다 관련 이벤트를 추출하고 처리하도록 한다. 시뮬레이터에서 사용되는 소셜 네트워크 그래프는 Facebook에서 추출된 소셜 네트워크 그래프[4]에서 300개의 노드를 샘플링 하여 구성하고, 각 노드는 평균 10개의 이웃 노드를 가지도록 설정하였다. 각 노드들 간의 에지 타입은 EMAIL, OSN, SMS, PHONE, MOUTH로 총 5가지로 설정하였고, 각각의 메시지 전송 지연 시간(T_{delay}^{type})을 나타내는 틱값은 EMAIL은 200, OSN은 150, SMS는 130, PHONE은 180로 설정하였다. MOUTH와 같은 경우는 만약 사용자가 일정 거리 내에 존재한다면 메시지 전송 지연 시간을 나타내는 틱값을 1로 하고 그렇지 않을 경우엔 거리에 비례해서 증가하도록 설정하였다. 또한 다양한 사용자 유형을 고려하여 표1과 같이 각 사용자 유형에 대한 채널 선호도와 재전송 확률을 정의하였다. 사용자의 유형은 기본적으로 균등하게 분포하도록 설정되어 있지만, 지역적 특성을 고려한 시나리오에서는 각 사용자의 위치에 따라 편향된 사용자 분포를 가질 수 있도록 하였다. 구현된 시뮬레이터의 인터페이스는 그림 5와 같다.

먼저 사용자 유형 분포가 메시지 전송 특성에 미치는 영향을 분석하기 위해 사용자들의 지역을 sparse하게 분포 시키고, 서로 다른 사용자 유형 분포를 적용하였을 때의 메시지 전송 범위와 지연 시간 분포를 측정하였다. 그림 6에서는 이 시나리오의

User Type	Preferred Channel	P_{fwd}
십대 (Teenagers)	OSN, SMS	mid
대학생 (CS Students)	OSN	low
가정주부 (Housewife)	PHONE, OSN	high
시니어 (Seniors)	MOUTH, PHONE	low
직장인 (Businessmen)	PHONE, SMS	high

표 1. 사용자 유형과 해당 선호채널 및 재전송 확률정도

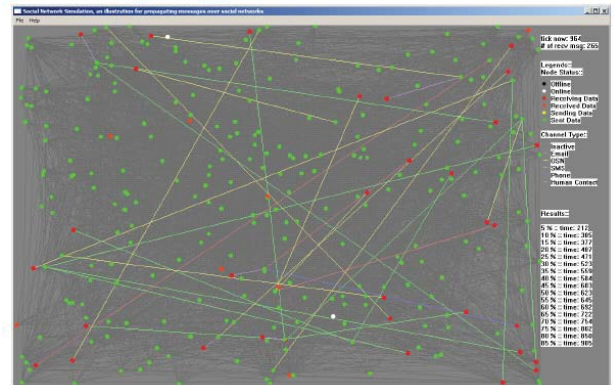


그림 5. 사용자 분포와 유형 설정 인터페이스

결과인 메시지 전송 지연 시간의 분포를 나타내고 있다. Uniform User Type Distribute는 사용자 유형을 결정할 때 모든 유형의 사용자들을 분포시킨 결과를 나타낸다. Business City(상업지구)는 직장인 사용자 유형이 선택될 확률을 아주 높게 설정한 결과를, 그리고 College City(대학지구)는 시니어 사용자 유형이 선택될 확률을 아주 높게 설정한 결과를 나타낸다. 그 결과를 살펴보면, 직장인들과 같이 긴밀한 연락을 주고받는 사용자들이 많을 경우 다양한 유형의 사용자들이 균등하게 분포되어 있는 경우보다 더욱 빠른 시간 안에 보다 많은 사용자에게 메시지가 전달되는 것을 알 수 있었다. 반면, 메시지 재전송 확률도 낮고 지연시간이 상대적으로 긴 채널을 선호하는 시니어 사용자 유형이 많은 경우에는 일반적인 균등한 사용자 유형 분포보다 메시지 전달시간이 확연하게 증가하는 것을 볼 수 있다. 또한 메시지 전송 범위도 약간 줄어드는 것을 확인 할 수 있었다.

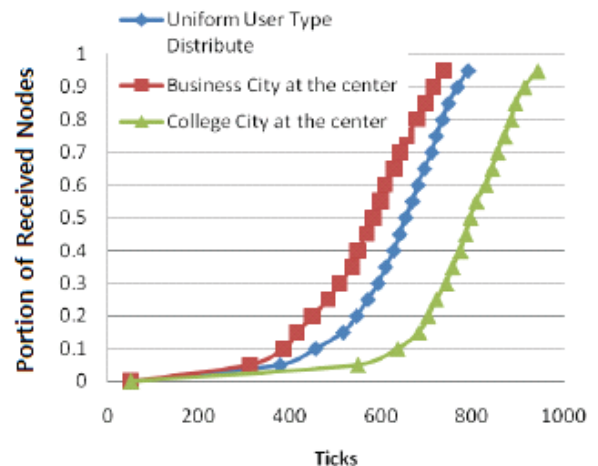


그림 6. 다양한 사용자 유형 분포에 따른 메시지 전송 결과
다음으로는 사용자의 지역 분포가 메시지 전송 특성에 미치는 영향을 살펴보았다. 그림 7에서는 대부분의 사용자들의 유형

이 직장인인 경우 사용자의 지역 분포가 메시지 전송 지연 시간에 미치는 영향에 대한 결과를 나타내고 있다. Urban Map은 사용자의 지역 분포가 urban분포를 따르는 경우에 대한 결과를 Uniform Node Distributed는 sparse분포에 대한 결과를 나타내고 있다. 결과적으로 직장인과 같이 메시지 전송에 적극적인 사용자 유형이 많더라도 사용자의 지역 분포가 메시지 전송 특성에 많은 영향을 주는 것을 확인 할 수 있었다. 이에 대한 주원인으로 MOUTH채널의 활성화를 생각할 수 있다. 즉 사용자간의 거리가 충분히 가깝다면 MOUTH채널을 통해 빠르게 메시지를 전달하는 것이 가능해져서 지연시간이 줄어드는 효과를 볼 수 있다.

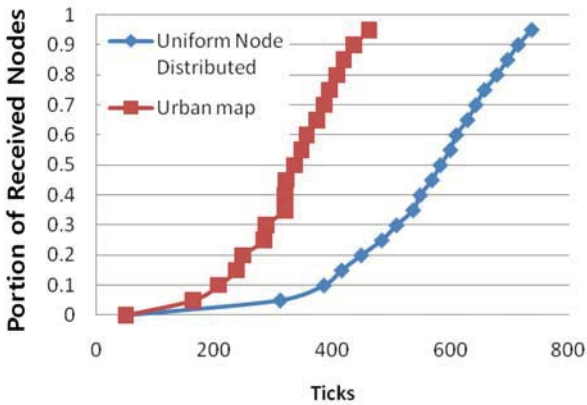


그림 7. 사용자 지역 분포에 따른 메시지 전송 결과

사용자의 지역 분포와 MOUTH 채널의 효과를 보다 자세히 확인하기 위해, 균등한 사용자 유형 분포를 가정 후 다양한 사용자 지역 분포를 적용하였을 때의 메시지 전송 결과를 분석하였다. 이를 위해 sparse, urban, two cities, three cities 사용자 지역 분포를 사용할 때 사용자간의 거리를 고려하지 않은 MOUTH 채널을 사용한 경우와 사용자간의 거리를 고려한 MOUTH 채널을 사용할 경우의 메시지 전송 지연 시간의 결과를 그림 8과 그림 9에서 나타낸다.

그림 8에서, 사용자간의 거리를 고려하지 않은 MOUTH채널을 사용한 경우에는 sparse 경우를 제외한 다른 사용자 지역 분포의 경우 메시지 전송 성능이 비슷함을 알 수 있었다. 이는 소셜 네트워크상의 사용자 지역 분포가 한곳 또는 여러 곳에 집중될수록 메시지 전송 성능이 향상됨을 의미한다. 반면, 사용자간의 거리를 고려한 MOUTH채널이 사용될 경우, 그림 9에서와 같이 사용자의 지역 분포가 밀집되는 구간의 개수에 따라서 그 성능이 달라지는 것을 확인 할 수 있었다. 사용자의 지역 밀집도가 가장 높은 urban 경우에 가장 짧은 평균 메시지 전송 지연 시간이 걸렸고, two cities, three cities, sparse 순으로 평균 메시지 전송 지연 시간이 증가함을 알 수 있다. two cities와 같은 경우 하나의 밀집된 구역의 노드가 다른 구역에 있는 이웃

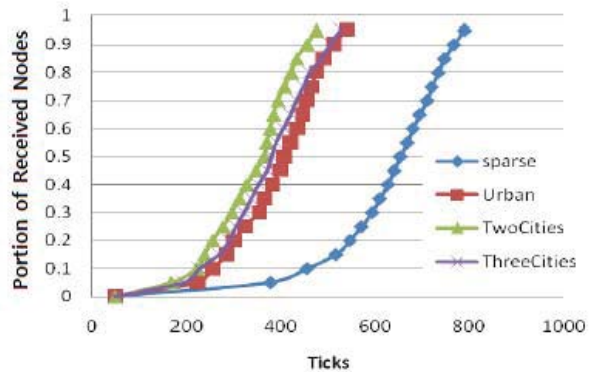


그림 8. 사용자간 거리를 고려하지 않은 MOUTH채널 사용 시 다양한 사용자 지역 분포에 따른 메시지 전송 결과

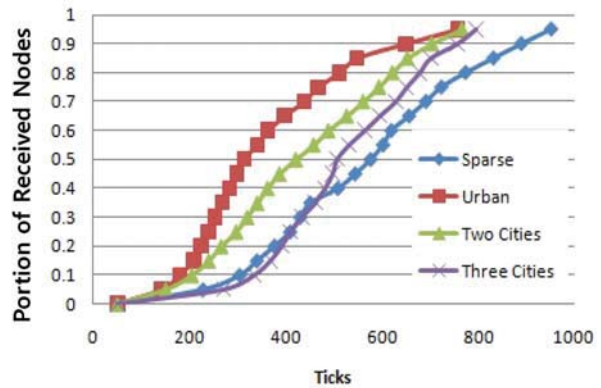


그림 9. 사용자간 거리를 고려한 MOUTH채널 사용 시 다양한 사용자 지역 분포에 따른 메시지 전송 결과

노드와 MOUTH채널을 통해 통신을 시도할 경우, 메시지 전송 지연시간이 크게 늘어나게 되고 전체 네트워크상의 평균 메시지 전송 지연 시간도 늘어나게 된다. 이처럼 구두전달 채널을 고려할 경우 예기치 못한 메시지 전송특성을 보일 수 있다는 점을 확인 할 수 있었고, 이 부분에 대한 연구가 앞으로 필요할 것으로 예상된다.

VI. 결론

온라인 소셜 네트워크의 대중화와 함께 실제 소셜 네트워크 정보를 인터넷에서 수집 할 수 있게 되었다. 이러한 인터넷 서비스에 내포된 소셜 네트워크 정보를 이용해 메시지 전송에 대한 연구를 수행함으로써, 정보 추천 시스템 및 경보 시스템 등에 응용할 수 있다. 이 논문에서는 다중 채널을 이용하는 소셜 네트워크상에서 이루어지는 메시지 전송 특성을 이해하기 위한 네트워크 모델(Delay Weighted Multi Channel Independent Cascade Model)과 이 모델을 기반으로 수행되는 메시지 전송 알고리즘을 제안하였다. 제안된 모델은 메시지 전송의 범위에 초점을 둔 기존의 모델을 확장하여 다양한 채널의 메시지 지연

시간, 사용자의 채널 선호도, 사용자의 유형 및 지역 분포를 고려함으로써 다양한 상황에서의 메시지 전송 범위 및 지연 시간을 분석할 수 있도록 하였다. 특히 구두전달 채널의 경우 그 사용여부가 메시지 특성에 예기치 못한 영향을 끼칠 수 있음이 확인 되었고 앞으로 이 부분에 연구를 진행할 예정이다.

REFERENCES

- [1] Michael Sirivianos, Kyungbaek Kim, Jian Wei Gan and Xiaowei Yang. Assessing the Veracity of Identity Assertions via OSNs. In Proceedings of the Fourth International Conference on COMMunication Systems and NETworkS (COMSNETS 2012), January 3-7, 2012, Bangalore, India
- [2] Michael Sirivianos, Kyungbaek Kim and Xiaowei Yang. SocialFilter: Introducing Social Trust to Collaborative Spam Mitigation. In Proceedings of the 30th IEEE International Conference on Computer Communications (IEEE INFOCOM 2011), April 10-15, 2011, Shanghai, China
- [3] Kyungbaek Kim, Ye Zhao and Nalini Venkatasubramanian. GSFord: Towards a Reliable Geo-Social Notification System. In Proceedings of the 31st IEEE International Symposium on Reliable Distributed Systems (SRDS 2012), October 8-11, 2012, Irvine, California, USA
- [4] Kyungbaek Kim. Sybil-Resistant Trust Value of Social Network Graph. In Proceedings of the First International Conference on Smart Media and Applications (SMA 2012), August 21-24, 2012, Kunming, Yunnan, China
- [5] Haewoon Kwak, Changhyun Lee, Hosung Park, Sue Moon. What is Twitter, a Social Network or a News Media? In Proceedings of WWW conference, 2010
- [6] Ceren Budak, Divyakant Agrawal and Amr El Abbadi. Limiting the Spread of Misinformation in Social Networks. In Proceedings of WWW conference, 2011
- [7] Wei Chen, Yajun Wang, Siyu Yang, Efficient Influence Maximization in Social Networks, In Proceedins of KDD 2009

저자 소개



김 경 백(정회원)

1999년 2월 한국과학기술원 전자
전산 학사 졸업.

2001년 2월 한국과학기술원 전자
전산 석사 졸업

2007년 2월 한국과학기술원 전자
전산 박사 졸업

2007년~2011년 University of California, Irvine, 박사
후연구원

2012년~현재 전남대학교 전자컴퓨터공학부 조교수

<주 관심분야 : 분산시스템, 미들웨어, 피어투피어 네
트워크, 오버레이 네트워크, 소셜 네트워크, 모
바일 클라우드 시스템.>