

대규모 워크플로우 소속성 네트워크를 위한 근접 중심도 랭킹 알고리즘[☆]

An Estimated Closeness Centrality Ranking Algorithm for Large-Scale Workflow Affiliation Networks

이 도 경¹ 안 현² 김 광 훈^{2*}
Do-kyong Lee Hyun Ahn Kwang-hoon Pio Kim

요 약

워크플로우 소속성 네트워크는 워크플로우 기반 조직의 수행자와 업무 사이의 연관관계를 나타내는 소셜 네트워크의 한 형태이며, 이를 기반으로 연결 중심도, 근접 중심도, 사이 중심도, 위세 중심도 등과 같은 다양한 분석 기법들이 제안되었다. 특히, 전사적 워크플로우 모델을 기반으로 하는 소속성 네트워크의 근접 중심도 분석은 워크플로우 소속성 네트워크의 규모가 증가함에 따라, 중심도 및 랭킹 계산의 시간 복잡도 문제점을 가진다는 것을 발견하였다. 본 논문에서는 근접 중심도 분석의 시간 복잡도 문제를 개선하기 위해, 근사치 추정 방법을 이용한 워크플로우 기반 소속성 네트워크의 추정 근접 중심도 기반 랭킹 알고리즘을 제안한다. 노드의 타입이 수행자인 워크플로우 예제 모델을 추정 근접 중심도 기반 랭킹 알고리즘에 적용한 성능 분석을 실시하였다. 수행 결과, 네트워크 규모 관점에서의 정확도는 평균적으로 47.5% 향상되었고, 샘플 모집단 비율 관점에서는 평균적으로 9.44% 정도의 향상된 수치를 보였다. 또한, 추정 근접 중심도 랭킹 알고리즘의 평균 계산 시간은 네트워크의 노드 수가 2400개, 샘플 모집단의 비율이 30%일 때, 기존 근접 중심도 랭킹 알고리즘의 평균 계산 시간보다 82.40%의 높은 성능을 보였다.

☞ 주제어 : 소속성 네트워크, 추정 근접 중심도, 랭킹 알고리즘

ABSTRACT

A type of workflow affiliation network is one of the specialized social network types, which represents the associative relation between actors and activities. There are many methods on a workflow affiliation network measuring centralities such as degree centrality, closeness centrality, betweenness centrality, eigenvector centrality. In particular, we are interested in the closeness centrality measurements on a workflow affiliation network discovered from enterprise workflow models, and we know that the time complexity problem is raised according to increasing the size of the workflow affiliation network. This paper proposes an estimated ranking algorithm and analyzes the accuracy and average computation time of the proposed algorithm. As a result, we show that the accuracy improves 47.5%, 29.44% in the sizes of network and the rates of samples, respectively. Also the estimated ranking algorithm's average computation time improves more than 82.40%, comparison with the original algorithm, when the network size is 2400, sampling rate is 30%.

☞ keyword : affiliation network, estimated closeness centrality, ranking algorithm

1. 서 론

최근 기업에서는 경쟁 산업 내에서 전략적 지위를 잡기 위해, 조직간, 기업 간의 비즈니스 프로세스 기반 협업이

증가하고 있으며, 이로 인해 워크플로우 기반 소셜 네트워크의 규모 또한 증가하고 있다. 이에 따라 워크플로우 기반 네트워크의 업무를 기반으로 한 인적 자원을 효율적으로 관리하기 위해서 워크플로우 관리시스템을 사용한다. 워크플로우 관리시스템은 비즈니스 프로세스, 인적 자원, 관련 데이터, 응용 프로그램 등을 자동으로 관리해주는 시스템이다. 워크플로우 관리시스템에 의해 생성된 워크플로우 기반 소셜 네트워크를 통해 수행자 중심의 인적 자원에 관련된 정보를 얻을 수 있다. 워크플로우 기반 소속성 네트워크는 워크플로우 기반 소셜 네트워크의 한 형태로서, 인적 자원 그룹의 업무 패턴이나 협업 관계 등에 대한 정보를 파악할 수 있다. 소속성 네트워크는 수행자의 집단과, 업무 집단으

¹ Management Information System, Kyonggi Univ., Iui-dong, Yeongtong-gu, Suwon-si, Gyeonggi-do, 443-760, Korea

² Computer Science, Kyonggi Univ., Iui-dong, Yeongtong-gu, Suwon-si, Gyeonggi-do, 443-760, Korea

* Corresponding author (kwang@kyonggi.ac.kr)

[Received 3 February 2015, Reviewed 4 February 2015(R2 30 June 2015), Accepted 14 October 2015]

☆ 본 연구는 한국연구재단에서 지원하는 2012년도 기초연구사업(No. 2012006971)의 연구수행으로 인한 결과물임.

로 나눌 수 있어, two-mode network라고도 불린다. 소속성 네트워크는 이분 그래프(bipartite graph)와 이분 행렬(bipartite matrix) 형태로 표현할 수 있으며, 소속성 네트워크의 이분 그래프를 통해서 특정 업무에 대한 공동 수행자 정보와 특정 수행자가 참여하고 있는 공동 업무 정보 등을 분석한다. 워크플로우 기반 조직의 규모의 증가로 인해 근접 중심도와 랭킹 계산 시간이 기하급수적으로 증가한다는 문제점을 인식하였다.

본 논문에서는 협업으로 인한 워크플로우 기반 소속성 네트워크의 규모가 증가됨에 따라, 기하급수적으로 늘어나는 소속성 네트워크의 인적 자원을 효율적으로 관리하고, 시간 복잡도 문제를 해결하기 위해 워크플로우 기반 소속성 네트워크에 추정 기법[5]을 적용한 추정 근접 중심도 기반 랭킹 알고리즘[4]을 제안하였다. 이를 통해 대규모 워크플로우 기반 소속성 네트워크의 랭킹을 구하는데 있어서 한계점을 해결한다. 또한, 추정 근접 중심도 기반 랭킹 알고리즘에 예제 모델을 적용해 제안한 알고리즘의 성능을 분석한다.

2. 관련 연구

21세기 정보화 시대의 정보 범람으로 인해, 전 세계는 무궁무진한 다양한 정보들을 효율적으로 관리하고, 신뢰도 있는 정확한 정보들을 얻기 위해 노력하고 있다. 이를 위한 방법으로 Page Rank가 제안되었으며, Google과 같은 검색 엔진들은 최적화된 정보를 제공해주고, 정보들을 효율적으로 관리하기 위해 노력하고 있다. Page Rank는 backlink를 기반으로 web graph를 통해 웹 페이지의 랭킹을 계산한다[14]. Backlink는 웹 페이지, 디렉토리, 웹 사이트, 도메인과 같은 여러 개의 웹 노드로부터 하나의 웹 노드가 받는 링크를 의미하며, 이를 통해, web graph가 만들어지고, web graph를 통해, 선호하는 웹 노드가 무엇인지, 그리고 중요한 웹 노드가 무엇인지 파악해, 웹 페이지들 간의 관계에 대한 중요성을 측정한다.

기업 간 그리고 조직 간의 협업의 증가와 글로벌화로 인해 워크플로우 기반 소셜 네트워크의 규모가 기하급수적으로 증가함에 따라 무한하게 생성되고 있는 자원들을 효율적으로 관리하기 위해 워크플로우 기반 조직 내의 인적 자원들 간의 소셜 네트워크를 분석하기 위한 워크플로우 모델과 이력 정보로부터 정보를 발견하는 워크플로우 마이닝[9], 워크플로우 발견 및 재발견 기술[11-13]과 같은 비즈니스 프로세스 인텔리전스에 대한 연구들이 선행되었다. 이후 워크플로우 기반 소셜 네트워크의 프레임 워크[10]를 제안하였고,

워크플로우 협력 네트워크[7]를 업무-수행자 소속성 네트워크로 재정의하고, 업무-수행자 소속성 네트워크의 분석을 위한 업무-수행자 이분행렬 생성 알고리즘을 제안하였다[2]. 소속성 네트워크의 분석 기법에는 연결 중심도(degree centrality), 근접 중심도(closeness centrality), 사이 중심도(betweenness centrality), 위세 중심도(eigenvector centrality) 등과 같은 다양한 분석 방법이 있다[3].

본 논문에서는 업무-수행자 이분행렬 생성 알고리즘을 기반으로 워크플로우 소속성 네트워크의 모델로부터 이분행렬을 생성하여, 업무수행자 소속성 네트워크의 근접 중심도를 구하고, 업무 혹은 수행자 집단의 노드 간의 관계를 분석한다. 더 나아가서, 워크플로우 기반 소속성 네트워크에 추정 기법인 RAND 알고리즘을 적용하여, 워크플로우 기반 소속성 네트워크의 규모가 커짐에 따라 발생하는 시간 복잡도 문제를 해결하기 위해 추정 근접 중심도 기반 랭킹 알고리즘을 제안한다.

3. 추정 근접 중심도 기반 랭킹 알고리즘

본 장에서는 전사적 워크플로우 기반 소속성 네트워크를 분석하기 위한 추정 근접 중심도 랭킹 알고리즘의 설계 및 수행과정에 대해 설명한다.

3.1 추정 근접 중심도 기반 랭킹 알고리즘의 설계 및 수행 과정

소속성 네트워크의 이분 그래프를 이분행렬로 나타내면, 이분 행렬($X^{P,A}$)은 다음의 표 1과 같다. 소속성 네트워크 내에서의 전체 수행자 수를 g , 전체 업무의 수를 h 라고 할 때, $X^{P,A}$ 의 크기는 $(g+h) \times (g+h)$ 이다. 소속성 네트워크

(표 1) 소속성 네트워크의 이분행렬
(Table 1) Bipartite graph of Affiliation Network

	φ_1	φ_2	φ_3	φ_4	φ_5	α_1	α_2	α_3
φ_1	0	0	0	0	0	1	1	0
φ_2	0	0	0	0	0	1	0	0
φ_3	0	0	0	0	0	0	1	0
φ_4	0	0	0	0	0	0	0	1
φ_5	0	0	0	0	0	0	1	1
α_1	1	1	0	0	0	0	0	0
α_2	1	0	0	0	1	0	0	0
α_3	0	0	0	1	1	0	0	0

의 이분행렬인 표 1은 업무 소속 관계를 나타내는 A 와 수행자 참여 관계를 나타내는 A' 를 포함하는 대칭 행렬이다. $X^{P,A}$ 의 부분 행렬인 A 와 A' 간의 연산을 통해 업무 협업 행렬(X^P)과, 수행자 공유 행렬(X^A)을 계산하여, 수행자 간의 업무 협업 관계와 업무 사이에서의 수행자 공유 관계를 분석한다.

$$X^P = A \times A' \quad (1)$$

$$X^A = A' \times A \quad (2)$$

(표 2) 업무 협업 행렬

(Table 2) Actor co-membership matrix

	φ_1	φ_2	φ_3	φ_4	φ_5
φ_1	2	1	1	0	1
φ_2	1	1	0	0	0
φ_3	1	0	1	0	1
φ_4	0	0	0	1	1
φ_5	1	0	1	1	2

(표 3) 수행자 공유 행렬

(Table 3) Event overlap matrix

	α_1	α_2	α_3
α_1	2	1	0
α_2	1	3	1
α_3	0	1	2

표 2의 업무 협업 행렬에서 φ_1 과 φ_5 는 두 개의 공동 업무를 수행하고 있다. 또, 표 3의 수행자 공유 행렬에서 업무 α_1 과 업무 α_3 은 수행자 2명이 참여하고 있고, 업무 α_2 에서는 수행자 3명이 공동으로 참여하고 있음을 알 수 있다. 즉, 협업 행렬과 수행자 공유 행렬의 각각의 대각 행렬(diagonal matrix)을 통해서 한 명의 수행자가 몇 개의 공동 업무에 참여하고 있는지 그리고, 하나의 업무에 몇 명의 수행자가 공동 참여하고 있는지에 대한 업무 협업 관계를 분석 할 수 있고, 이를 통해 워크플로우 기반 소속성 네트워크의 인적 자원을 효율적으로 관리할 수 있다.

추정 근접 중심도 기반 랭킹 알고리즘[5]을 구현하기에 앞서, 워크플로우 기반 소속성 네트워크를 추정 근접 중심도 기반 랭킹 알고리즘에 적용시키기 위해서 수행자와 업무 노드의 연관 관계에 따라 표 1과 같이 특정 수행자 노드와 특정 업무의 노드가 소속되어 있다면 1값, 그렇지 않으면 0값을 갖는 이분 행렬의 형태로 표현한다.

$$x_{i,j} = \begin{cases} 1 & \text{수행자}(\varphi_i) \text{와 업무}(\alpha_j) \text{가 소속되어 있는 경우} \\ 0 & \text{소속되어 있지 않은 경우} \end{cases} \quad (3)$$

이후, 대규모 워크플로우 기반 소속성 네트워크에서의 노드 간의 최단거리를 구하기 위해서 모든 정점 사이의 최단 경로를 한 번에 찾아 주는 Floyd 알고리즘[17]을 사용한다. Floyd 알고리즘에 들어갈 가중치 행렬(weighted matrix)의 초기 값은 이분 행렬을 바탕으로 대각 행렬은 모두 0값을 갖으며, 1과 대각 행렬을 제외한 0값은 가중치 행렬에서 무한값을 갖는다. Floyd 알고리즘의 결과로 생성된 최단거리 행렬(shortest-path matrix)을 통해 추정 근접 중심도와 기존 근접 중심도를 구하고, 랭킹알고리즘을 통해 순위를 구하고 그 결과를 비교 분석한다.

Ranking Algorithm of Closeness Centralities for WSAN: RankCCWSAN(t, k)

Input A Performer-Activity Affiliation Network Model, $G(\varphi, \alpha)$
 /* 수행자-업무 소속성 네트워크 모델 */
 Output Ranking list /* 랭킹 리스트 */

- t 는 순위 매길 정점의 타입 (수행자 or 업무)
- k 는 순위 매길 정점의 개수
- P 는 수행자의 집합
- A 는 업무의 집합
- S 는 샘플 정점들의 집합. $S = \{s_1, \dots, s_l\}$
- l 은 샘플 정점의 개수
- \hat{C} 는 추정된 근접 중심도의 집합
- C 는 측정된 근접 중심도의 집합
- g 는 전체 업무 수행자의 수
- h 는 전체 업무의 수

Begin Procedure

```

Initialize  $\hat{C} = \emptyset, C = \emptyset$ 
If ( $t = \text{"PERFORMER"}$ ) /* 순위 매길 정점의 타입이 수행자인 경우 */
     $S \leftarrow \text{RANDOM}(P, l)$ ;
    For ( $\forall \varphi_i \in P$ ) Do
        Begin
            /* 수행자 정점들의 추정된 근접 중심도 구하기 */
             $\hat{C}_{\varphi_i} = \left[ \sum_{j=1}^g \frac{g d(S_j, \varphi_i)}{l(g-1)} \right]^{-1}$ ;
             $\hat{C} = \hat{C} \cup \hat{C}_i$ ;
        End
    Sort  $\hat{C}$ ; /* 추정된 근접 중심도를 내림차순으로 정렬 */
    For ( $\hat{C}_i \in \hat{C}; i \leq k$ ) Do
        Begin
            /* 기존 알고리즘을 통해 근접 중심도 구하기 */
             $C_{\varphi_i} = \left[ \sum_{\varphi_j} d(\varphi_i, \varphi_j) \right]^{-1} (\varphi_i \neq \varphi_j)$ ;
             $C = C \cup C_i$ ;
        End
    Sort  $C$ ;
    Return  $C$ ;
else /* 순위 매길 정점의 타입이 업무인 경우 */
     $S \leftarrow \text{RANDOM}(A, l)$ ;
    For ( $\forall \alpha_i \in A$ ) Do
        Begin
            /* 업무 정점들의 추정된 근접 중심도 구하기 */
             $\hat{C}_{\alpha_i} = \left[ \sum_{j=1}^h \frac{h d(S_j, \alpha_i)}{l(h-1)} \right]^{-1}$ ;
             $\hat{C} = \hat{C} \cup \hat{C}_i$ ;
        End
    Sort  $\hat{C}$ ; /* 추정된 근접 중심도를 내림차순으로 정렬 */
    For ( $\hat{C}_i \in \hat{C}; i \leq k$ ) Do
        Begin
            /* 기존 알고리즘을 통해 근접 중심도 구하기 */
             $C_{\alpha_i} = \left[ \sum_{\alpha_j} d(\alpha_i, \alpha_j) \right]^{-1} (\alpha_i \neq \alpha_j)$ ;
             $C = C \cup C_i$ ;
        End
    Sort  $C$ ;
    Return  $C$ ;
End Procedure
    
```

(그림 1) 추정 근접 중심도 랭킹 알고리즘

(Figure 1) Estimated closeness centrality ranking algorithm

소속성 네트워크는 수행자와 업무 두 가지 속성을 가지고 있기 때문에, 두 가지 관점으로 나누어 추정 근접 중심도 기반 랭킹 알고리즘에 적용할 때, 수행자 혹은 업무집단에 대한 가치 있는 정보를 얻을 수 있다. 그림 1은 근접 중심도 추정 알고리즘과 기존의 근접 중심도 분석 알고리즘[1]을 결합하여 랭킹을 계산하는 추정 근접 중심도 랭킹 알고리즘이다.

추정 근접 중심도 랭킹 알고리즘은 순위 매길 정점의 타입(t)과, 샘플링 할 정점의 개수(l)를 RAND알고리즘에 입력 하면서 시작된다. 순위 매길 정점의 타입을 수행자(φ)라고 할 때, 샘플 수행자 집합에 속하는 정점들(s)은 추정 알고리즘과 기존 알고리즘을 통해 각각의 추정 근접 중심도와 기존 근접 중심도를 구하게 되며, 내림차순 정렬을 통해 랭킹 값을 구한다. 정렬된 추정 근접 중심도의 랭킹 값과 기존의 측정된 근접 중심도의 랭킹 값을 결과 값으로 돌려받고, 추정 근접 중심도 분석 랭킹 알고리즘은 종료된다.

본 논문에서 추정 근접 중심도 랭킹 알고리즘에 적용한 워크플로우 기반 소속성 네트워크 모델은 임의로 만든 예제 모델이며, 워크플로우 기반 소속성 네트워크의 규모를 부풀릴 때, 임의로 만든 예제 모델을 확장하게 되는데, 이 때, 확장하는 배수만큼, 같은 value 값을 갖는 노드들이 존재하게 되며, 알고리즘의 랭킹 결과 값 또한, 같은 경우가 발생하게 된다. 추정 근접 중심도 랭킹 알고리즘의 랭킹 결과 값은 알고리즘의 Input 값으로 받는 워크플로우 기반 소속성 네트워크에 따라서 달라지기 때문에, 임의의 모델이 아닌 실제 워크플로우 모델을 적용했을 때에는 지금과는 다른 결과를 얻을 수 있다. 표 4는 추정 근접 중심도 기반 랭킹 알고리즘의 결과 값으로 수행자 모델의 노드 개수가 12개인 임의의 example model을 5배로 확장하여, 총 수행자 노드의 개수가 60개, 순위 매길 노드의 개수를 30개, 샘플 모집단의 비율을 10%로 지정하고, 수행했을 때의 결과이다.

(표 4) 추정 근접 중심도 랭킹 알고리즘 랭킹 결과
(Table 4) Ranking result of estimated closeness centrality

Performer Size=60, Rank Size=30 Sampling rate=10%		
node id	랭킹 결과	추정 근접 중심도 결과 값
node ₉ , node ₂₁ , node ₃₃ , node ₄₅ , node ₅₇	1	0.2
node ₀ , node ₁₂ , node ₂₄ , node ₃₆ , node ₄₈	6	0.1
node ₁₁ , node ₂₃ , node ₃₅ , node ₄₇ , node ₅₉	11	0.06666667

node ₁₀ , node ₂₂ , node ₃₄ , node ₄₆ , node ₅₈	15	0.022222223
node ₅ , node ₁₇ , node ₂₉ , node ₄₁ , node ₅₃	21	0.011111111
node ₃ , node ₄ , node ₈ , node ₁₅ , node ₁₆	26	0.008695652

표 4에서는 node₉, node₂₁, node₃₃, node₄₅, node₅₇ 가 0.2의 추정 근접 중심도 값을 갖으며, 1위의 랭킹 결과를 얻었다.

3.2 근접 중심도의 추정 기법

소속성 네트워크의 근접 중심도는 네트워크 안에서 한 노드와 다른 노드들 간의 최단거리(SSSP-Single Source Shortest Path)[3]의 합을 의미한다. 워크플로우 기반 소속성 네트워크의 규모가 커짐에 따라 기존 근접 중심도를 구하는 기법[3]대로, 소속성 네트워크의 모든 노드와 모든 노드 사이의 최단 거리를 구하는 데에는 한계가 있다. 따라서 전체 소속성 네트워크에서 샘플 정점들을 랜덤하게 추출해, 샘플 정점들과 모든 노드 사이의 추정된 근접 중심도를 구함으로써, 시간을 단축할 수 있는 RAND알고리즘을 적용시킨다. 아래의 수식 (4)는 추정 근접 중심도를 구하는 수식이며, 추정된 근접 중심도 수식에서의 $d(S_j, \varphi_i)$ 는 랜덤하게 뽑아낸 샘플 정점들과 소속성 네트워크 내에서의 모든 정점들 간의 최단거리를 의미한다.

$$\widehat{C}_{\varphi_i} = \left[\sum_{j=1}^l \frac{gd(S_j, \varphi_i)}{l(g-1)} \right]^{-1} \quad (4)$$

RAND알고리즘을 통해 도출된 추정 근접 중심도는 기존 근접 중심도 분석 알고리즘에서의 근접 중심도를 구하는 방법과 비교하였을 때, 근접 중심도를 구하는 시간을 단축시킬 수 있어, 대규모 워크플로우 기반 소속성 네트워크의 규모에 따른 문제점을 해결 할 수 있다.

4. 추정 근접 중심도 기반 랭킹 알고리즘 성능분석

본 논문에서는 추정 근접 중심도 기반 랭킹 알고리즘의 성능을 분석하기 위하여, 워크플로우 기반 소속성 네트워크의 예제 모델을 적용하여 알고리즘을 구현하였다. 예제 모델의 노드 타입은 수행자(φ)이며, 순위 매길 노드의 개수(k)는 20개로 모두 동일한 조건에서, 워크플로우 기반 소속성 네트워크의 규모와 임의로 선택한 샘플 모집단의 비율 관점에서 정확도 분석을 하였으며, 추정근접 중심도 기반 랭킹 알고리즘

과 기존 근접 중심도 랭킹 알고리즘의 평균 계산 시간을 비교 분석하였다.

아래 표 5는 추정 근접 중심도 기반 랭킹 알고리즘의 성능을 분석하기 위한 실험 환경 내에서의 파라미터에 대한 설명이다. RankCCSWAN은 추정 근접 중심도 랭킹 알고리즘을 의미하며, Naive는 기존 근접 중심도 랭킹 알고리즘을 의미한다. M은 워크플로우 기반 소속성 네트워크의 규모이며, S는 워크플로우 기반 소속성 네트워크에서 임의로 선택된 샘플 모집단의 비율을 의미한다. m_{avg} 는 소속성 네트워크의 규모 관점에서 분석한 정확도의 평균값이며, s_{avg} 는 소속성 네트워크에서 임의로 선택된 모집단의 크기 관점에서 분석한 정확도의 평균값이다.

(표 5) 실험 매개변수의 정의
(Table 5) Test parameter definition

parameter	정 의
RankCCSWAN	추정 근접 중심도 랭킹 알고리즘
Naive	기존 근접 중심도 랭킹 알고리즘
M	추정 근접 중심도 랭킹 알고리즘에서 소속성 네트워크의 규모
S	추정 근접 중심도 랭킹 알고리즘에서 임의로 선택된 샘플 모집단의 비율
m_{avg}	소속성 네트워크의 규모 관점에서 분석한 정확도의 평균 값
s_{avg}	소속성 네트워크에서 임의로 선택된 샘플 점들의 규모 관점에서 분석한 정확도의 평균 값

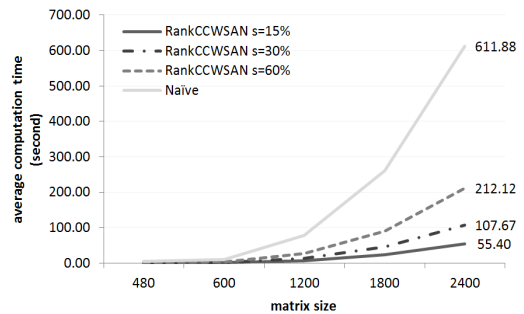
표 6은 추정 근접 중심도 랭킹 알고리즘의 정확도 분석 결과이다. 정확도를 분석하는 기준은, 추정 근접 중심도의 랭킹 결과 값과, 기존 근접 중심도 랭킹 결과 값의 동일한 정도이다.

(표 6) 추정 근접 중심도 랭킹 알고리즘 정확도 분석결과
(Table 6) Result of accuracy analysis of estimated closeness centrality ranking algorithm

M \ S	s				m_{avg}
	s=15%	s=30%	s=60%	s=100%	
60	25	20	65	100	52.5
120	0	35	90	100	56.25
240	20	40	100	100	65
360	90	80	100	100	92.5
480	100	100	100	100	100
600	100	100	100	100	100
1200	100	100	100	100	100
1800	100	100	100	100	100
2400	100	100	100	100	100
s_{avg}	70.56	75	95	100	

추정 근접 중심도 랭킹 알고리즘의 소속성 네트워크 규모 관점에서 살펴보면, M이 60일 때, 정확도의 평균값은 52.5%인 반면, M이 2400일 때, 정확도의 평균값은 100%로, 47.5%의 높은 정확도를 보였다. 워크플로우 기반 소속성 네트워크에서 임의로 선택된 샘플 모집단의 비율 관점에서 살펴보면, S가 15%일 때, 정확도의 평균값은 70.56%, S가 100%일 때, 정확도 평균값이 100%로, 정확도 평균값이 29.44% 증가했다. 즉, 워크플로우 기반 소속성 네트워크의 규모가 커질수록, 임의로 선정해내는 샘플 모집단의 규모가 커질수록 높은 정확도를 갖는 분석 결과를 얻었다.

그림 2는 추정 근접 중심도 랭킹 알고리즘과 기존 근접 중심도 랭킹 알고리즘의 평균 계산 시간을 비교한 것이다.



(그림 2) 평균 계산 시간 비교
(Figure 2) Comparison of the average computation time

M이 2400일 때, RankCCSWAN의 샘플 모집단의 비율이 15%, 30%, 60%일 때, 각각의 평균 계산 시간은 55.40초, 107.67초, 212.12초로 Naive의 평균 계산 시간인 611.88초 보다 각각 90.95%, 82.40%, 65.33%의 향상된 계산 시간을 보였다.

5. 결 론

본 논문에서는 대규모 워크플로우 기반 소속성 네트워크의 규모 확대에 의해 발생하는, 방대한 양의 업무 패턴, 업무와 수행자 정보 등의 인적 자원 관리의 한계를 해결하기 위해서, 추정 방법을 적용한 추정 근접 중심도 기반 랭킹 알고리즘을 제안하였다. 또한 추정 근접 중심도 기반 랭킹 알고리즘의 성능을 분석하기 위해 정확도와 평균 계산 시간을 분석해 보았다. 그 결과, 워크플로우 기반 소속성 네트워크의 규모와, 랜덤하게 추출해내는 샘플 모집단의 비율이 커질수록, 높은 정확도를 얻을 수 있다는 결론을 얻었으며, 추정 근접 중심도 랭킹 알고리즘의 평균 계산 시간은 기존 근접 중심도 랭킹 알

고리즘의 평균 계산 시간보다 평균적으로 78.83%정도 향상된 평균 계산 시간을 보였다. 향후 연구에서는 추정 근접 중심도 랭킹 알고리즘의 정확성을 높이기 위해서 추정 과정에서 샘플링 모집단을 지능적으로 선택하기 위한 알고리즘 및 방법에 대하여 연구하고자 한다.

참 고 문 헌 (Reference)

- [1] A. Battestseg, H. Ahn, Y. Lee, M. Park, H. Kim, W. Yoon, K. Kim, "Organizational Closeness Centrality Analysis on Workflow-supported Activity-Performer Affiliation Networks", Proceedings of the 15th International Conference on Advanced Communication Technology, pp. 154-157, 2013. http://ieeexplore.ieee.org/xpls/abs_all.jsp?amumber=6488160
- [2] H. Ahn, K. Kim, "An Activity-Performer Bipartite Matrix Generation Algorithm for Analyzing Workflow-supported Human-Resource Affiliations", Journal of Korean Society for Internet Information, Vol. 14, No. 2, pp. 25-34, 2013. <http://dx.doi.org/10.7472/jksii.2013.14.2.25>
- [3] D. Knoke, S. Yang, SOCIAL NETWORK ANALYSIS - 2nd Edition, Series : Quantitative Applications in the Social Science, SAGE Publications, 2008.
- [4] D. Lee, H. Ahn, H. Kim, K. Kim, "A Ranking Algorithm of Closeness Centralities for Large-Scale Workflow-supported Affiliation Networks", Korean Society for Internet Information, Vol. 15, No. 1, pp. 37-38. 2014.
- [5] D. Lee, H. Ahn, H. Kim, K. Kim, "Accuracy Analysis of an Estimated Closeness Centrality based Ranking Algorithm in Large Scale Workflow-supported Affiliation Networks", Korean Society for Internet Information, Vol. 15, No. 2, pp. 171-172, 2014.
- [6] A. Aini, A. Salehipour, "Speeding up the Floyd - Warshall algorithm for the cycled shortest path problem.", Applied Mathematics Letters Vol. 25, No. 1, pp. 1-5, 2012. <http://dx.doi.org/10.1016/j.aml.2011.06.008>
- [7] H. Kim, H. Ahn, K. Kim, "A Workflow Affiliation Network Discovery Algorithm", ICIC Express Letters(SCOPUS), Vol. 6, No. 3, pp. 765-770, 2012.
- [8] J. Song, M. Kim, H. Kim, K. Kim, "A Framework : Workflow-based Social Network Discovery and Analysis", Proceedings of the 13th International Conference on Computational Science and Engineering, pp. 421-426, 2010. <http://dx.doi.org/10.1109/CSE.2010.74>
- [9] Van Der Aalst, W. M, H. A. Reijers, and M. Song., "Discovering social networks from event logs." Journal of Computer Supported Cooperative Work (CSCW), Vol. 14, No. 6, pp. 549-593, 2005. <http://dx.doi.org/10.1007/s10606-005-9005-9>
- [10] J. Song, M. Kim, H. Kim, K. Kim, "A Framework : Workflow-based Social Network Discovery and Analysis", Proceedings of the 13th International Conference on Computational Science and Engineering, pp. 421-426, 2010.
- [11] M. Park, K. Kim, "Control-Path Oriented Workflow Intelligence Analysis", Journal of Information Science and Engineering, Vol. 24, No. 2, pp. 343-359, 2008. <http://dx.doi.org/10.1109/ICCT.2007.212>
- [12] M. Park, K. Kim, "A Workflow Event Logging Mechanism and Its Implications on Quality of Workflows", Journal of Information Science and Engineering, Vol. 26, No. 5, pp. 1817-1830, 2010.
- [13] K. Kim, C.A. ELLIS, "Workflow Reduction for Reachable-path Rediscovery in Workflow Mining", Foundations and Novel Approaches in Data Mining, Vol. 9, pp. 288-309, 2006. http://dx.doi.org/10.1007/11539827_17
- [14] Page. Lawrence, et al., "The PageRank citation ranking: Bringing order to the web.", 1999. <http://ilpubs.stanford.edu:8090/422/>
- [15] K. Okamoto, W. Chen, X.Y. Li, "Ranking of Closeness Centrality for Large-Scale Social Networks", Lecture Notes in Computer Science of Frontiers in Algorithmics, Vol. 50 No. 59, pp. 186-195, 2008.
- [16] K. Faust, "Centrality in affiliation networks", Social Networks, Vol. 19, No. 2, pp.157-191, 1997.
- [17] Borgatti, Stephen P., and Daniel S. Halgin, "Analyzing affiliation networks." The Sage handbook of social network analysis, pp. 417-433, 2011.
- [18] H. Ahn, C. Park, K. Kim, "A correspondence analysis framework for workflow-supported performer-activity affiliation networks", Advanced Communication Technology (ICACT), 2014 16th International Conference on. IEEE, 2014.
- [19] E. Yan , D. Ying, "Applying centrality measures to impact analysis: A coauthorship network analysis." Journal of the American Society for Information Science and Technology, Vol. 60, No. 10, pp. 2107-2118, 2009.
- [20] T. Frank, "A Family of Affiliation Indices for Two-Mode Networks." Journal of Social Structure Vol. 14, 2013.

● 저 자 소개 ●



이 도 경 (Do kyong Lee)

2012.3~현재 경기대학교 경상대학 경영정보학과 학사(재학)
관심분야 : 워크플로우/비피엠, Workflow-supported Affiliation Networks
E-mail : gaiety93@kgu.ac.kr



안 현 (Hyun Ahn)

2011년 경기대학교 컴퓨터과학과(학사)
2013년 경기대학교 일반대학원 컴퓨터과학과(석사)
2013년~현재 경기대학교 일반대학원 컴퓨터과학과(박사과정)
관심분야 : 워크플로우/비피엠, 비즈니스프로세스 인텔리전스
E-mail : hanh@kgu.ac.kr



김 광 훈 (Kwang hoon Kim)

1984.2 경기대학교 전자계산학과(학사)
1986.2 중앙대학교 전자계산학과(석사)
1986.2~1991.8 한국전자통신연구원 연구원
1994.5 University of Colorado Boulder, Department of Computer Science, MS
1998.5 University of Colorado Boulder, Department of Computer Science, Ph.D.
2005.3~2010.2 Univ. of Colorado Boulder, Department of Computer Science, 방문교수
2007.7~2010.6 경기대학교 콘텐츠융합소프트웨어연구센터장
1998.3~현재 경기대학교 컴퓨터과학과 교수
2000.1~현재 한국인터넷정보학회 이사, 부회장
2002.3~현재 비피엠코리아포럼 부회장
2003.1~현재 WiMC ERC Vice-chair
2003.1~현재 TTA 정보통신국제표준전문가
관심분야 : 워크플로우/비피엠, Process-Aware Information Systems, Process Discovery/Rediscovery,
Workflow-supported Social Networks
E-mail : kwang@kyonggi.ac.kr