

제약 조건을 갖는 버블정렬 그래프의 연결도 분석

서정현* · 이형옥**

*순천대학교 로봇과학교육센터, **순천대학교 컴퓨터교육과

Analysis of Bubblesort graph's connectivity which has a conditions for limitations

Joungh-hung Seo* · Hyeong-ok Lee** (교신저자)

Sunchon National Univ.

E-mail : {jhseo, oklee}@scnu.ac.kr

요 약

버블정렬 그래프는 자료를 정렬하는 버블정렬 기법을 수학적 모델인 그래프로 모델링한 것이다. 버블정렬 그래프 B_n 는 분지수 n , 라우팅 경로 길이 $\frac{n(n-1)}{2}$, 망비용 $O(n^3)$ 이다. 본 연구에서는 버블정렬 그래프의 망비용을 개선하기 위한 방법으로 버블정렬 그래프의 분지수 개수를 대략 $\frac{1}{2}$ 정도 줄인 제약조건을 제시한다. 이러한 제약 조건을 갖는 버블정렬 그래프는 임의의 노드 U 에서 V 까지 라우팅을 위해 연결된 그래프임을 증명한다.

ABSTRACT

Bubblesort graph is mathematically modeled with bubbling methods, which can arrange data. Bubblesort graph B_n 's degree is n , its routing path length $\frac{n(n-1)}{2}$, and its network cost is $O(n^3)$. In this paper we suggest the number of Bubblesort graph's degree reduced to half as a solution to improve the network cost of Bubblesort graph. The Bubblesort graph which has the following restriction is a connected graph randomly from node U to node V for routing.

키워드

버블정렬 그래프, 분지수, 라우팅, 연결된 그래프

I. 서 론

그래프 이론(Graph Theory)은 수학의 조합론(Combinatorial theory)의 일부로서 고전적인 학문 분야이다. 그래프 이론이 현실문제에 본격적으로 활용되기 시작한 것은 2차 세계대전으로 탄생한 OR(Operation Research)의 영향이 크다. 오늘날에도 산업공학에서 max-flow를 활용한 최적화 계산 모형은 현실적 활용의 사례이고, 슈퍼컴퓨팅의 빠른 발전과 더불어 그래프 이론은 전산학 전 분야에서 다양하게 활용되고 있으며 현대 수학에서 중요한 이론적 도구가 되고 있다. 다양한 컴퓨터 분야 중에서 인류 문명의 획기적인 변화를 가져올 분야로 인공지능이 꼽히고 있다. 기존의 인공지능 분야는 인간적인 사고를 신경망 이론, 카오스 이론, 퍼지 이론 등으로 실현하려고 했지만 발전이 더디게 진행되고 있다. 근

래 바둑 분야에서 알파고의 등장은 4차 산업혁명의 이슈를 몰고 왔고, 인공지능 분야에서 「데이터 기반 학습 후 추론」이 강력한 대체 수단으로 부상했다. 「데이터 기반 학습 후 추론」은 대용량의 데이터와 빠른 연산이 필수적으로 요구 된다.

슈퍼컴퓨터는 지금까지 분자 생물학, 일기예보, 양자역학, 비행 시뮬레이션, 핵융합, 우주 탐사 등 매우 많은 계산을 요구하는 분야에서 빠른 속도로 계산하기 위해 사용되었다. 2016년 11월 발표된 세계 최고 성능의 컴퓨터 순위에서 1, 2위는 중국의 선웨이 타이후라이트와 텐허-2가 차지하였고, 3~5위는 미국, 6~7위는 일본이다. 중국의 선웨이 타이후라이트의 처리속도는 $93 \times 10^{15} flops$ (floating point operations per second)이다. 일본은 2020년까지 $10 \times 10^{15} flops$ 성능의 슈퍼컴퓨터

터를 개발할 예정이다.

고성능 병렬처리 컴퓨터 시스템의 기반 기술로 활용할 수 있는 상호연결망은 그래프 문제로 모델링하여 그래프 이론 관점에서 평가척도(분지수, 대칭성, 확장성, 이분할 성질, 사이클, 연결도, 그래프 채색 등)를 분석하고 최단경로, 사이클, 고장허용 라우팅, 트리 구조, 정렬, 방송, 병렬 경로 등의 알고리즘을 통해 다양한 학문 분야에서 응용될 수 있는 기반을 제공한다.

지금까지 널리 알려진 상호연결망은 메쉬 부류(torus, hexagonal, honeycomb, cube-connected cycle 등), 하이퍼큐브 부류(hypercube[4], folded hypercube, de bruijin, butterfly, hierarchical cubic, 재귀원형군, 하이퍼-스타 등), 스타 그래프 부류(star[1], pancake[2], bubblesort[6], bubblesort star, star-mobius, bubblesort star, transposition, macro-star, matrix-star, (n,k)-star, generalized-star cube, star-crossed cube 등) 등이 제안되었다.

n개의 원소로 노드 주소를 표현하는 버블정렬 그래프 B_n 의 노드는 $\langle N \rangle = \{1, 2, 3, \dots, n\}$ 의 순열로 노드 주소를 나타낸다. 버블정렬 그래프의 에지는 순열의 인접한 2개 심볼이 서로 교환된 순열 사이에 인접하다. 따라서 노드 개수는 $n!$ 이고, 분지수는 $(n-1)!$ 이고, 그래프 전체 에지 개수는 $n!(n-1)/2$ 개다. 버블정렬 그래프 B_n 의 노드 B의 주소가 $B = b_1b_2b_3 \dots b_{j+1} \dots b_n$ 이라 하자. 노드 B와 에지 $k_j, 1 \leq j \leq n-1$ 에 의해 인접한 노드의 순열은 심볼 b_j 와 b_{j+1} 이 서로 교환된 $b_1b_2b_3 \dots b_{j+1}b_j \dots b_n$ 이다.

버블정렬 그래프 B_n 의 임의의 노드 U에서 V까지 라우팅은 노드 U의 순열에서 n개 심볼을 오름차순으로 정렬하는 것과 동일하게 생각할 수 있다. 따라서 버블정렬 그래프의 지름은 $n(n-1)/2$ 이다. 버블정렬 그래프는 노드 주소에서 가장 오른쪽 위치한 주소 심볼을 기준으로 n개의 부분 그래프로 분할할 수 있으며, 부분 그래프는 재귀적으로 (n-1)차원 버블정렬 그래프와 동일한 구조를 갖는다. 버블정렬 그래프는 스타 그래프와 동일한 분지수를 갖지만 지름이 큰 값을 가지므로 망비용이 $O(n^3)$ 을 갖는다.

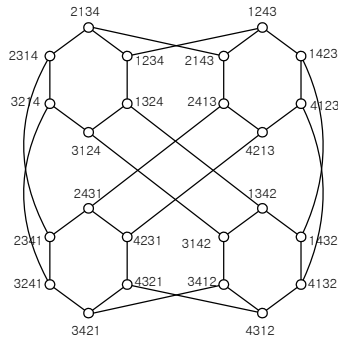


그림 1. 4차원 버블정렬 그래프

본 연구에서는 버블정렬 그래프의 망비용(=지름 × 분지수)인 $O(n^3)$ 을 개선하기 위해 제약조건을 갖는 버블정렬 그래프에 대해 연구한다. 망비용을 개선하는 방법은 첫째, 분지수를 줄이는 방법 둘째, 지름을 개선하는 방법으로 생각해 볼 수 있다. 본 연구에서는 망비용을 개선하기 위한 초기 단계로 분지수를 개선하는 방법을 먼저 적용한다.

II. 본 론

1. 버블정렬 그래프의 제약 조건

n-차원 버블정렬 그래프 B_n 는 n개 심볼을 이용하여 노드 주소 $B = b_1b_2b_3 \dots b_{j+1} \dots b_n$ 로 표현한다. 그래프 B_n 에지는 노드 주소에서 인접한 심볼들 사이에 에지가 존재하므로 (n-1)개의 에지가 존재한다. 본 연구에서는 에지 개수를 줄이면서도 그래프가 연결된 상태(connected graph)를 가지도록 제약 조건을 제시한다.

기존 버블정렬 그래프에서 에지에 대한 제약 조건에 의해 생성된 에지는 에지 $B_f, 1 \leq f \leq \lfloor \frac{n}{2} \rfloor$ 과 에지 B_i 이 있다($n \geq 3$).

(1) 에지 $B_f, 1 \leq f \leq \lfloor \frac{n}{2} \rfloor$

노드 B의 주소를 $B = b_1b_2b_3 \dots b_{j+1} \dots b_n$ 이라 할 때, 노드 B와 에지 B_f 에 인접한 노드 $B_f(B)$ 주소는 $b_f(B) = b_1b_2b_3 \dots b_{f+1}b_f \dots b_n$ 이다.

(2) 에지 B_i

노드 B와 에지 B_n 에 인접한 노드 $B_i(B)$ 의 주소는 $B_i(B) = b_n \dots b_{f+1}b_f \dots b_3b_2b_1$ 이다.

따라서 제약 조건을 갖는 버블정렬 그래프 B_n 의 노드 개수는 $n!$ 개이고, 노드 B에 인접한 에지 $B_f, 1 \leq f \leq \lfloor \frac{n}{2} \rfloor$ 의 개수는 $\lfloor \frac{n}{2} \rfloor$ 이고, 에지 B_n 의 개수는 1개이므로 분지수는 $\lfloor \frac{n}{2} \rfloor + 1$ 이다. 결과적으로 버블정렬 그래프의 에지 개수를 대략 1/2로 줄인 결과를 갖는다.

2. 제약조건을 갖는 버블정렬 그래프 연결도 분석

그래프의 연결도는 노드 연결도와 에지 연결도로 나눈다. 그래프에서 에지를 제거하는 것은 에지에 부속한 노드를 제거하는 것을 의미하므로 본 연구에서는 에지를 기준으로 연결도를 분석하여 버블정렬 그래프에서 제약조건을 갖는 경우에도 연결된 그래프임을 보인다.

정리 1 제약조건을 갖는 버블정렬 그래프 B_n 은 연결된 그래프(connected graph)이다.

증명 제약조건을 갖는 버블정렬 그래프의 노드 S의 주소를 $s_1s_2s_3 \dots s_{i-1}s_i s_{i+1} \dots s_{j-1}s_j s_{j+1} \dots s_n$ 이라 하자($i \neq j, i < j$). 제약조건을 갖는 버블정렬 그래프

(본 논문의 이하 내용에서는 간단히 BB_n 로 표현)가 연결된 그래프라면 임의의 노드 S에서 노드 T까지 경로가 존재한다. 임의의 노드 S에서 T까지 경로가 존재한다는 것은 BB_n 그래프에서 노드 S에서 $(n-1)$ 개의 서로 다른 노드의 순열을 생성할 수 있음을 의미하고, 노드 S에서 노드 T까지 라우팅을 위한 경로상의 노드가 제약조건을 갖는 버블정렬 그래프의 에지 정의에 의해 생성 가능함을 의미한다.

본 증명에서는 BB_n 의 노드 S의 주소 $S = s_1s_2s_3 \dots s_i \dots s_j \dots s_n$ 에서 제약조건을 갖는 에지 정의에 의해 심볼 s_i 와 심볼 s_j 가 임의의 위치로 교환 가능함을 통해 $n!$ 개의 노드 주소를 생성 가능함을 보인다. 노드 S의 주소 $s_1s_2s_3 \dots s_i \dots s_j \dots s_n$ 에서 심볼 s_i 와 심볼 s_j 가 위치한 조건에 따라 3가지로 나누어 증명한다.

경우1. $(i \neq j) \leq \lfloor \frac{n+1}{2} \rfloor, i < j$

BB_n 의 에지 B_j 는 $B_j = \{B_1, B_2, B_3, \dots, B_{\lfloor \frac{n}{2} \rfloor}\}$ 이다. 노드 S의 주소 $S = s_1s_2s_3 \dots s_{i-1}s_i s_{i+1} \dots s_{j-1}s_j s_{j+1} \dots s_n$ 에서 $(i \neq j) \leq \lfloor \frac{n+1}{2} \rfloor$ 인 경우 인덱스 j 의 최대 값은 $\lfloor \frac{n+1}{2} \rfloor$ 이다. 심볼 $s_{j = \lfloor \frac{n+1}{2} \rfloor}$ 는 BB_n 의 에지 B_j 에 의해 심볼 $s_{\lfloor \frac{n}{2} \rfloor}$ 와 교환 가능하다. BB_n 의 에지 $B_{\lfloor \frac{n}{2} \rfloor}$ 에 의해 인덱스 i 의 최대 값은 $\lfloor \frac{n}{2} \rfloor$ 이다.

노드 S의 심볼 s_i 와 심볼 s_j 를 교환하기 위해 심볼 s_i 와 s_j 는 서로 인접한 위치에 있어야 한다. 노드 S에서 심볼 s_i 를 이동하여 심볼 s_j 와 인접하도록 하기 위해 에지시퀀스 $K' = \langle B_i, B_{i+1}, B_{i+2}, \dots, B_{j-3}, B_{j-2} \rangle$ 를 적용한다. 노드 S에서 에지시퀀스 K' 를 적용한 노드의 순열은 $K'(S) = s_1s_2s_3 \dots s_{i-1}s_i s_{i+1} \dots s_{j-1}s_j s_{j+1} \dots s_n$ 이다. 노드 $K'(S) = s_1s_2s_3 \dots s_{i-1}s_i s_{i+1} \dots s_{j-1}s_j s_{j+1} \dots s_n$ 에서 인접한 2개 심볼 s_i 와 s_j 를 교환하는 에지는 B_{j-1} 이다. 노드 $K'(S)$ 에서 에지 B_{j-1} 에 인접한 노드는 $B_{j-1}(K'(S))$ 이고 노드의 순열은 $B_{j-1}(K'(S)) = s_1s_2s_3 \dots s_{i-1}s_i s_{i+1} \dots s_{j-1}s_j s_{j+1} \dots s_n$ 이다. 논문의 편의상 노드 $B_{j-1}(K'(S))$ 를 노드 S' 라 하자. 즉 노드 $B_{j-1}(K'(S)) = S'$ 이다. 노드 $S' = s_1s_2s_3 \dots s_{i-1}s_i s_{i+1} \dots s_{j-1}s_j s_{j+1} \dots s_n$ 에서 심볼 s_j 를 왼쪽에서 i 번째 위치하도록 이동하려면 에지시퀀스 $K'' = \langle B_{j-2}, B_{j-3}, \dots, B_{i+2}, B_{i+1}, B_i \rangle$ 를 적용하면

된다. 노드 S' 에서 에지시퀀스 $K'' = \langle B_{j-2}, B_{j-3}, \dots, B_{i+2}, B_{i+1}, B_i \rangle$ 를 적용한 노드 $K''(S') = s_1s_2s_3 \dots s_{i-1}s_i s_{i+1} \dots s_{j-1}s_j s_{j+1} \dots s_n$ 이다. 따라서 노드 S의 주소 $S = s_1s_2s_3 \dots s_{i-1}s_i s_{i+1} \dots s_{j-1}s_j s_{j+1} \dots s_n$ 에서 에지시퀀스 $\langle B_i, B_{i+1}, B_{i+2}, \dots, B_{j-3}, B_{j-2}, B_{j-1}, B_{j-2}, B_{j-3}, \dots, B_{i+2}, B_{i+1}, B_i \rangle$ 를 적용하면 심볼 s_i 와 s_j 는 서로 교환된다.

경우2. $(i \leq \lfloor \frac{n+1}{2} \rfloor)$ and $(j > \lfloor \frac{n+1}{2} \rfloor)$

노드 S의 심볼 s_i 와 s_j 를 교환하기 위해서는 에지 $B_j = \{B_1, B_2, B_3, \dots, B_{\lfloor \frac{n}{2} \rfloor}\}$ 을 이용해야 한다. 에지 B_j 를 이용하기 위한 심볼 s_i 와 s_j 의 인덱스 조건은 $|s_i| \leq \lfloor \frac{n+1}{2} \rfloor$ 이고 $|s_j| \leq \lfloor \frac{n+1}{2} \rfloor$ 이다. 노드 $S = s_1s_2s_3 \dots s_i \dots s_j \dots s_n$ 에서 심볼 s_i 는 에지시퀀스 $K' = \langle B_i, B_{i+1}, B_{i+2}, \dots, B_{\lfloor \frac{n}{2} \rfloor} \rangle$ 를 적용하여 심볼 s_i 의 인덱스 $|s_i| = \lfloor \frac{n+1}{2} \rfloor$ 위치로 이동한다. 심볼 s_j 는 에지 B_n 을 적용하면 전체 심볼이 역순으로 배열되므로 심볼 s_j 의 인덱스 $|s_j| < \lfloor \frac{n+1}{2} \rfloor$ 이 된다. 심볼 s_j 와 심볼 s_i 를 서로 인접하도록 적용할 에지시퀀스는 HB_n 에서 $n=2k+1$ (즉, 홀수)과 $n=2k$ (즉, 짝수)로 나누어진다.

BB_n 에서 $n=2k+1$ (즉, 홀수)인 경우 에지시퀀스 $K^{odd} = \langle B_{n-j+1}, B_{n-j+2}, B_{n-j+3}, \dots, B_{\lfloor \frac{n}{2} \rfloor - 1} \rangle$ 를 적용하고, $n=2k$ (즉, 짝수)인 경우 에지시퀀스 $K^{even} = \langle B_{n-j+1}, B_{n-j+2}, B_{n-j+3}, \dots, B_{\lfloor \frac{n}{2} \rfloor - 2} \rangle$ 를 적용하면 심볼 s_j 와 심볼 s_i 는 서로 인접한다. 노드 S의 인접한 심볼 s_j 와 심볼 s_i 을 교환하는 에지는 $n=2k+1$ (즉, 홀수)인 경우 에지 $B_{\lfloor \frac{n}{2} \rfloor}$ 을 적용하고, $n=2k$ (즉, 짝수)인 경우 에지 $B_{\lfloor \frac{n}{2} \rfloor - 1}$ 을 적용한다. 노드 S의 심볼 s_i 와 심볼 s_j 를 교환한 후에는 가장 마지막에 적용한 에지 $B_{\lfloor \frac{n}{2} \rfloor}$ (또는 $B_{\lfloor \frac{n}{2} \rfloor - 1}$)을 제외하고 지금까지 적용한 에지시퀀스를 역순으로 적용하여 나머지 순열의 원소를 제자리로 되돌린다.

$K' = \langle B_i, B_{i+1}, B_{i+2}, \dots, B_{\lfloor \frac{n}{2} \rfloor} \rangle,$

$$K^{odd} = \langle B_{n-j+1}, B_{n-j+2}, B_{n-j+3}, \dots, B_{\lfloor \frac{n}{2} \rfloor - 1} \rangle,$$

$$K^{even} = \langle B_{n-j+1}, B_{n-j+2}, B_{n-j+3}, \dots, B_{\lfloor \frac{n}{2} \rfloor - 2} \rangle$$

$n=2k+1$ (즉, 홀수)인 경우 적용할 전체 에지시퀀스 K_{2k+1} 는 다음과 같다.

$$K_{2k+1} = \langle K', B_n, K^{odd}, B_{\lfloor \frac{n}{2} \rfloor}, \overline{K^{odd}}, B_n, \overline{K'} \rangle$$

$n=2k$ (즉, 짝수)인 경우 적용할 전체 에지시퀀스 K_{2k} 는 다음과 같다.

$$K_{2k} = \langle K', B_n, K^{even}, B_{\lfloor \frac{n}{2} \rfloor}, \overline{K^{even}}, B_n, \overline{K'} \rangle$$

경우3. ($i \neq j$) $\left\lceil \frac{n+1}{2} \right\rceil$

노드 $S = s_1 s_2 s_3 \dots s_i \dots s_j \dots s_n$ 의 심볼 s_i 와 심볼 s_j 의 인덱스는 ($|s_i| \neq |s_j|$) $\left\lceil \frac{n+1}{2} \right\rceil$ 이므로 현재 상태에서 심볼 교환을 수행할 수 있는 에지 B_j 가 없다. 노드 S 에서 에지 B_n 을 적용하면 노드 $B_n(S) = s_n \dots s_j \dots s_i \dots s_3 s_2 s_1$ 이 된다. 노드 $B_n(S)$ 의 심볼 s_i 와 심볼 s_j 의 인덱스는 ($|s_i| \neq |s_j|$) $\left\lfloor \frac{n+1}{2} \right\rfloor$ 이 되어 경우1과 동일한 조건이 되므로 경우1을 적용하면 된다.

따라서 노드 S 에서 BB_n 의 에지 정의에 의해 임의의 노드 순열이 생성 가능하므로 BB_n 그래프는 연결된 그래프이다. □

III. 결 론

상호연결망은 분산 및 병렬처리 시스템의 통신망의 최대 지연 시간을 최소 비용으로 줄일 수 있어서 병렬처리 컴퓨터의 기반을 제공한다.

기존의 버블정렬 그래프는 각 노드에 분지수가 $(n-1)$ 이고 망비용이 $O(n^3)$ 이다. 본 연구에서는 기존의 버블정렬 그래프의 망비용을 개선하기 위해 그래프의 에지 정의를 제약조건을 갖도록 하여 분지수를 $1/2$ 로 줄인 새로운 방법을 제시하였다.

본 연구에서 제시한 제약조건을 갖는 버블정렬 그래프는 분지수가 기존의 버블정렬 그래프보다 $1/2$ 로 줄어들었지만 그래프 자체는 연결된 그래프 (connected graph)를 유지함을 알 수 있다. 분지수가 줄어들어도 그래프가 연결되어 있음은 임의의 정점에서 다른 노드로 라우팅이 가능함을 알 수 있다. 이후의 연구 과제로는 제약조건을 정형화하여 그래프 이론 관점에서 추가적인 연구가 필요하다.

참고문헌

[1] S. B. Akers, D. Harel, B. Krishnamurthy, "The star graph: An attractive to the n-cube," Proc. Int. Conf. Parallel Processing, pp. 393-400, 1987.

[2] P. Berthome, A. Ferreira and S. Perennes, "Optimal Information Dissemination in Star and Pancake Networks," IEEE Trans. on Parallel and Distributed Syst., vol. 7, no. 12, pp. 1292-1300, 1996.

[3] S. B. Akers and B. Krishnamurthy, "A Group-Theoretic Model for Symmetric Interconnection Networks," IEEE Trans. Comput., vol. 38, no. 4, pp. 555-565, 1989.

[4] K. Day and A. Tripathi, "A Comparative Study of Topological Properties of Hypercubes and star graphs," IEEE Trans. Parallel and Distributed Systems, Vol. 5, no. 1, pp. 31-38, 1994.

[5] T. Yuuki. "On the Stack Number and the Queue Number of the Bubble-Sort Graph." IEICE TRANSACTIONS on Fundamentals of Electronics, vol. 99, no. 6, pp. 1012-1018, 2016.

[6] Y. Suzuki, and K. Kaneko, "The Container Problem in Bubble-Sort Graphs," IEICE Trans. Information and Systems, Vol. E91-D, No. 4, pp. 1003-1009, 2008.