추가제약을 가진 MST문제를 위한 실용적 근사해법
(A pragmatic approximation algorithm for constrained minimum spanning tree problem.)

홍성필*, 민대현**

(*중앙대학교 상경학부, **중앙대학교 대학원 경영학과 석사과정)

초록

최근 Ravi와 Goemanns는 즉 전체 길이의 합이 일정한 값을 넘지 않는 최소비용신장나무(minimum spanning tree problem)를 구하는 문제의 (1+ε, 1)-근사해를 구할 수 있는 알고리듬을 제시하였다. 즉 비용은 최적을 보장하지만 전체길이 제약조건은 근사적으로 만족하는 해를 생성한다. 그러나 이러한 알고리듬은 문제의 비가능해를 생성할 수 있으며 1/ε에 대하여 지수함수의 수행시간을 갖는다. 본 논문에서는 Ravi와 Goemanns의 알고리듬을 실용적으로 변형하여 전체 길이 제약조건을 정확히 만족하며, 그 비용은 최적비용과의 차이가 호의 비용 중 최대값을 넘지 않도록 보장하는 강성다항식 알고리듬을 제시한다.

1. 문제정의 및 수리모형

$G=(N,E)$를 무방향네트워크, 그리고 $|N|=n$, $|E|=m$이라 하자. 최소비용신장나무문제(MST)는 G의 신장나무 중에서 각 호 $e\in E$에 주어진 비용 c_e 의 합 $c(T)=\sum_{e\in T}c_e$을 최소로 하는 나무 T를 구하는 문제로, 가장 손쉽게 풀 수 있는 네트워크최적화문제들 가운데 하나이다. 여기서 T가 만족해야하는 추가적인 제약을 고려해보자. 즉 각 호 $e\in E$에 또 하나의 계수("길이") $d_e\geq 0$을 주고, T의 총길이 $d(T)=\sum_{e\in T}d_e$가 어떤 $\Delta\geq 0$보다 작도록 해를 국한한다. 이 문제를 추가제약을 가진 최소비용신장나무문제(Constrained Minimum Spanning Tree Problem, CMST)라고 부르자.

단순한 MST문제와는 달리 CMST는 쉽게 최적해가 구해지는 문제는 아니다. 그러나 소위 "weakly NP-hard"로 어느 정도 실용적인 해법은 개발되었으며 [1], 더욱이 최근에는, 어떤 고정된 오차계수 $\epsilon>0$에 대해 문제크기의 다항식으로 주어지는 수행시간 안에, 총길이제약을 정확히 만족하면서, 그 비용은 최적값의 $1+\epsilon$을 넘지 않는 $(1,\ 1+\epsilon)$-근사해를 구할 수 있는 소위 "polynomial approximation scheme, PAS"이 연구되어 왔다.

지금까지 알려진 "최선"의 알고리듬은 Ravi와 Goemanns[2]의 연구로 $O(m\log^2 n+n\log^3 n)$시간 서브루틴을 $O(n^{O(1/\epsilon)})$회 반복하면, 총길이제약은 $1+\epsilon$의 오차로 만족하지만, 총비용은 CMST의 최적비용 이하가 되는, $(1+\epsilon,\ 1)$-근사해를 보장하는 확장된 의미의 FPAS(fully (strongly-)polynomial approximation scheme)를 제시하였다. 그러나 이 알고리듬은 무엇보다도 가능해를 보장하지 못한다는 단점이 있으며, 가능해를 보장하는 $(1,\ 1+\epsilon)$-근사알고리듬으로 전환할 수는 있지만 이 경우 알고리듬의 수행시간은 유사다항식이 될 수밖에 없는 단점이 있다.

우리는 이 논문에서 $O(m\log^2 n+n\log^3 n)$ 수행시간 안에 최적비용과의 차이가 어떤 호의 비용을 넘지 않는 가능해를 구하는 실용적 알고리듬을 제시하려고 한다. 이 알고리듬은 Ravi와 Goemanns의 FPAS의 한 번의 반복단계에 해당하는 부분을 실용적으로 변형한 것이다. 이 알고리듬은 라그랑지안 완화법에 기반하고 있으며, 다음의 CMST 수리모형에 기반하고 있다.

집합 S를 G의 신장나무를 나타내는 incidence vector들의 집합이라고 하자. 그러면 CMST의 정수계획 모형은 다음과 같이 쓸 수 있다.

$$z^* = \min\ c(T)$$

(CMST) $\quad$ s.t. $T\in G$ 의 신장나무집합

$$d(T)\leq \Delta \qquad (1)$$

CMST를 라그랑지 승수를 써서 완화한다면, 완화대상이 되는 제약식은 자연스럽게 (1)을 선택할 수 있으며 이 때 완화된 문제 LR_λ는 다음과 같이 쓸 수 있다.

$$(LR_\lambda)$$

$$z(\lambda)=\min\ z(\lambda,T)=c(T)-\lambda(\Delta-d(T))$$

s.t. $T\in G$ 의 신장나무 집합

잘 알려진 바와 같이 모든 $\lambda \geq 0$에 대하여 $z(\lambda)$은 최적해의 값 z^*의 하한값이된다. 또한 $z(\lambda)$는 λ의 "concave piece- wise linear" 함수가 되며 따라서 다음의 값을 가지는 λ^*가 존재하게 된다.

$$z_{LR} = \max\{z(\lambda): \lambda \geq 0\}.$$

2. 몇 가지 성질들

앞에서 정의된 것처럼 $z_{LR} = z(\lambda^*)$일 때, LR_{λ^*}의 최적해 신장나무의 집합을 $\mathcal{T}$라고 하자. 잘 알려진 바와 같이 $\varepsilon > 0$가 충분히 작으면 $LR_{\lambda^*-\varepsilon}$과 $LR_{\lambda^*+\varepsilon}$의 최적신장나무들은 모두 $\mathcal{T}$에 속하게 된다. T_-를 $LR_{\lambda^*-\varepsilon}$의 하나의 최적신장나무라고 하자.

$$z(\lambda^* - \varepsilon) = c(T_-) - (\lambda^* - \varepsilon)(\Delta - d(T_-))$$

$$\leq \; z(\lambda^*) = c(T_-) - (\lambda^*)(\Delta - d(T_-)).$$

$$\Rightarrow \; d(T_-) \geq \Delta$$

유사한 방법으로 $LR_{\lambda^*+\varepsilon}$의 임의의 최적신장나무 T_+는 $d(T_+) \leq \Delta$를 만족한다는 것을 보일 수 있다. 앞에서 언급한 것처럼 T_-와 T_+는 $\mathcal{T}$에 속하기 때문에 우리는 다음의 정리를 증명한 것이 된다.

정리 1. λ^*에 대응되는 신장나무들 $\mathcal{T}$중에는 전체길이 제약 Δ를 만족하는 신장나무와 만족하지 않는 신장나무들이 동시에 존재한다.

다음의 정리는 CMST의 최적비용 z^*와 $\mathcal{T}$의 신장나무들과의 관계를 규명하는데 중요한 역할을 하게된다.

정리 2. $\mathcal{T}$의 임의의 신장나무를 T^*라고 할 때, T^*는 다음 문제의 최적해가 된다.

$$\min \; c(T)$$

$$\text{s.t. } T \in G \text{ 의 신장나무집합}$$

$$d(T) \leq d(T^*)$$

증명: T를 위의 문제의 가능해 신장나무라고 하자. 그러면 다음의 관계가 성립한다.

$$c(T^*) - \lambda^*(\Delta - d(T^*))$$

$$\leq \; c(T) - \lambda^*(\Delta - d(T))$$

$$\Rightarrow \; c(T^*) - c(T) \leq \lambda^*(d(T) - d(T^*))$$

T가 위의 문제의 가능해이므로 $d(T) - d(T^*) \leq 0$이며 따라서 $c(T^*) \leq c(T)$가 성립하여 정리가 증명된다. ■

정리 2.에 의하여 만일 $d(T^*) \geq \Delta$이면 정리에서 정의된 문제의 해집합은 CMST의 해집합보다 크므로 $c(T^*) \leq z^*$가 성립한다.

따름정리 3. $\mathcal{T}$에는 두 종류의 신장나무들만이 존재하는데, 첫째는 $d(T^*) \geq \Delta$와 $c(T^*) \leq z^*$며, 두 번째 종류는 $d(T^*) \leq \Delta$와 $c(T^*) \geq z^*$를 만족한다.

3. 근사알고리듬

본 논문에서 제시하려고 하는 근사알고리듬은 MST문제의 다음과 같이 잘 알려진 성질을 사용한다.

정리 4. 신장나무 T와 T'가 신장나무 incidence vector polytope에서 이웃할(adjacent) 필요충분 조건은 T가 T'로 부터 호 한 개씩을 서로 교환하여 얻어질 수 있는 것이다.

예를 들어 $\mathcal{T}$의 신장나무들은 선형계획문제 LR_{λ^*}의 polytope과 $c + \lambda^* d$로 정의되는 hyperplane과 만나는 optimal facet상의 vertex들에 해당하며 이들은 호의 교환에 의하여 모두 생성할 수 있다.

Ravi와 Goemanns의 (1+ε, 1)-근사알고리듬은 정리 2.에 의하여 존재성이 증명된 $d(T_+) \leq \Delta$을 만족하는 $T_+ \in \mathcal{T}$에서 시작하여 $d(T_-) \geq \Delta$를 만족하는 $T_- \in \mathcal{T}$까지의 신장나무 sequence를 생성하여 근사가능해를 구하는 것을 요지로 한다. 이에 비하여 우리의 알고리듬은 따름정리 3.을 사용하여 $d(T_-) \geq \Delta$ ($c(T_-) \leq z^*$)에서 $d(T_+) \leq \Delta$ ($c(T_+) \geq z^*$)로의 sequence를 생성하여 가능해생성에 중점을 둔다. 즉 T_-의 호들을 $T_+ \setminus T_-$의 호와 교환하면서 최초로 가능해를 구했을 때, 그 신장나무를 근사해로 채택한다. 따름정리 3.에 의해 최초로 가능해가 생성되는 순간이, 최초로 신장나무의 비용이 최적비용 z^*보다 같거나 커지는 순간이며 그 때 z^*와의 차이는 교환되어 들어오는 호의 비용을 넘지 않을 것이다.

정리 5. 근사알고리듬의 해의 비용을 z_{ALG}라고 할 때 $z_{ALG} \leq z^* + \max_{e \in E} c_e$.

알고리듬을 완성하기 위해서는 $\lambda^{*}, T_{-}, T_{+} \in \mathfrak{I}$를 구하는 방법과 그 수행시간을 밝혀야한다. 그러나 이 방법들은 앞에서 언급한 바와 같이 Ravi와 Goemanns의 알고리듬[2]과 동일하며 $O(m\log^{2}n + n\log^{3}n)$시간 안에 수행될 수 있다.

참고문헌

[1] Aggarwal, V., Y. Aneja and K. Nair, "Minimal spanning tree subject to a side constraint", *Comput. Operations Res*, 9, 287-296, 1982.

[2] Ravi, R. and M. X. Goemanns," The Constrained Minimum Spanning Tree Problem, manuscript, 1996.