2005 한국경영과학회/대한산업공학회 춘계공동학술대회
2005년 5월13일~14일, 충북대학교

고립된 노드가 없는 최대가중치 k-노드 부분그래프

명 영 수
단국대학교 천안캠퍼스 경상학부

초록: 그래프가 주어져 있고 각 노드에 가중치가 주어져 있다고 가정하자. 두 노드가 에지로 연결되어 있는 경우에 두 노드는 인접한다고 정의한다. 주어진 그래프의 부분그래프 중에서 k개의 노드로 이루어지고 부분그래프의 각 노드별로 하나 이상의 인접하는 노드가 부분그래프에 존재하면 이러한 부분그래프를 고립된 노드가 없는 k-노드 부분그래프라고 부른다. 본 논문에서는 고립된 노드가 없는 k-노드 부분그래프 중에서 부분그래프에 포함된 노드들의 가중치의 합이 최대가 되는 부분그래프를 선택하는 문제를 다룬다. 이 문제는 CDMA를 기반으로 하는 멀티캐스트 네트워크 설계에 응용 예를 갖고 있다. 본 논문에서는 대상문제의 복잡성을 규명하고, 상한과 하한을 구하는 해법을 개발한다.

1. 서론

고립된 노드가 없는 최대가중치 k-노드 부분그래프는 다음과 같이 정의된다. 노드(node)와 에지(edge)로 구성된 무방향 네트워크(undirected network)와 각 노드별 가중치가 주어져 있다고 가정하자. 두 노드가 에지로 연결되어 있는 경우에 두 노드는 인접한다고 정의한다. 그래프에서 에지가 연결되어 있지 않은 노드를 고립된 노드(isolated node)라고 부른다. 즉 고립된 노드는 인접하는 노드를 갖고 있지 않게 된다. 주어진 그래프의 부분그래프 중에서 부분그래프의 각 노드별로 하나 이상의 인접하는 노드가 부분그래프에 포함되는 경우에 이러한 부분그래프를 고립된 노드가 없는 부분그래프라고 부르기로 하자. 본 논문에서는 고립된 노드가 없는 k-노드 부분그래프 중에서 부분그래프에 포함된 노드들의 가중치의 합이 최대가 되는 부분그래프를 선택하는 문제를 다룬다. 이 문제는 CDMA를 기반으로 하는 멀티캐스트 네트워크 설계에 응용 예를 갖고 있다.

CDMA 네트워크에서도 멀티캐스트 서비스가 구현되고 있으며 멀티캐스트 서비스의 표준화가 중요한 관심사로 대두하고 있다 [1,2,3]. CDMA 네트워크에서의 멀티캐스트의 속성은 유선네트워크에서의 멀티캐스트와 크게 다르지는 않다. 즉 하나의 링크에 하나의 데이터를 흘려보낸다는 멀티캐스트의 근본적인 속성은 마찬가지이다. 그러나 무선통신의 특성상 사용자들은 전파를 통해서 동일한 데이터를 수신한 뒤에 각 사용자가 필요한 데이터를 추출함으로서 자신에게 특정된 서비스를 제공받는다는 점에서 각 사용자가 별도의 경로에 의해서 데이터를 수신하는 유선네트워크와 차별화된다. 이러한 차이 때문에 CDMA 네트워크에서의 각 사용자는 복수의 기지국에서 데이터를 동시에 수신할 수 있게 되고, 수신된 복수의 데이터를 융합함으로써 획기적으로 수신에러를 감소시킬 수 있다. 이러한 데이터의 융합기술을 soft-combine이라고 부른다.

CDMA 네트워크에서 soft-combine이 가능하려면 인접한 두 기지국에서 송출하는 데이터의 시간대를 공조시켜야 되는데, 이를 위해서는 Base Station Controller (BSC) 또는 Radio Network Controller (RNC)와 같은 네트워크 controller가 controller와 기지국간의 경로에 대한 정보를 유지하고 지연(delay) 효과 등을 파악하는 등 많은 추가적인 기능을 수행하여야 한다. 따라서 네트워크 controller의 용량 때문에 모든 기지국에 soft-combine이 가능하도록 지원할 수는 없고 정해진 수의 기지국에 대해서만 지원이 가능하다. 물론 soft-combine의 특성상 어떤 기지국이 선정되면 이 기지국과 인접한 기지국이 하나 이상은 선정되어야 soft-combine의 기능이 가능하다. 따라서 이러한 조건을 만족하면서 soft-combine의 효과를 극대화 하도록 soft-combine을 지원할 기지국을 선정하는 것이 CDMA를 기반으로 하는 멀티캐스트 네트워크 설계의 중요한 이슈이다.

위에서 언급한 멀티캐스트 네트워크 설계문제는 고립된 노드가 없는 최대가중치 k-노드 부분그래프를 구하는 문제로 표현할 수 있다. 기지국에 노드를 대응시키고, 두 노드에 대응되는 기지국이 인접하면 두 노드를 에지로 연결한다. 기지국이 관할하는 영역, 즉 셀(cell)에 속한 트래픽의 수요를 노드의 가중치로, soft-combine을 지원할 수 있는 최대 기지국의 수를 k 라고 가정하면, 멀티캐스트 네트워크 설계문제는 고립된 노드가 없는 최대가중치 k-노드 부분그래프를 구하는 문제가 된다.

2. 용어의 정의 및 문제의 복잡성

주어진 네트워크에서 노드의 집합을 $V=\{1,\cdots,n\}$, 무방향 에지의 집합을 $E=\{1,\cdots,m\}$로 표시한다. 두 노드 $i\in V$와 $j\in V$에 걸쳐있는 에지를 $e=(i,j)$로 표시하기로 하고 두 노드 i와 j를 에지 e의 종단노드(end nodes)라고 부르기로 한다. 주어진 그래프 G에는 두 노드를 동일한 종단노드로 하는 복수의 에지(multiple edge)나 동일한 노드를 종단노드로 하는 에지(self-loop)는 존재하지 않는 것으로 가정한다. 각 노드 $i\in V$의 가중치를 $d(i)$로 표시하기로 한다. 노드 i에 인접한 노드의 집합을 $N(i)$로 표시하기로 하자.

고립된 노드가 없는 최대가중치 k-노드 부분그래프를 구하는 문제는 전형적인 조합최적화문제의 형태를 취하고 있으나 기존에 알려진 문제는 아니다. 따라서 해당문제가 NP-hard 문제인지 여부를 판정하는 것이

우선적인 관심사이다. 만약에 이 문제가 NP-hard 문제라면 다항시간 내에 문제를 해결할 수 있는 해법은 기대하기 어려울 것이며, 이와 반대의 경우라면 그러한 해법을 발견하는 것이 중요한 과제가 될 것이기 때문이다.

정리 1. 고립된 노드가 없는 최대가중치 k-노드 부분그래프를 구하는 문제는 NP-hard이다.

(증명) 노드-커버(node cover)문제를 이용하여 증명 가능하나 자세한 증명은 생략한다. □

따름정리 2. 주어진 그래프가 planar 그래프인 경우에도 고립된 노드가 없는 최대가중치 k-노드 부분그래프를 구하는 문제는 NP-hard이다.

(증명) planar 그래프에 정의된 노드-커버 문제를 이용하여 동일한 방법으로 증명 가능. □

3. 상한의 도출

이 장에서는 대상문제를 표현하는 0-1 정수계획 모형을 수립하고 수립된 모형의 선형계획법 완화 문제를 이용하여 목적함수의 하한을 구하는 방법에 대해서 소개한다.

3.1 수리계획법 모형

고립된 노드가 없는 최대가중치 k-노드 부분그래프를 구하는 문제는 다음과 같이 정수계획 모형으로 표현할 수 있다. 노드 i가 선택되면 1, 아니면 0의 값을 갖는 변수 $x(i)$를 정의하자. 그러면 대상문제는 다음과 같이 표현이 가능하다.

$$(IP)\quad \max\ z=\sum_{i\in V} d(i)x(i) \qquad (1)$$

$$s.t.\ \sum_{i\in N(i)} x(i)\le k, \qquad (2)$$

$$x(i)-\sum_{j\in N(i)} x(j)\le 0, \qquad \forall i\in V, \qquad (3)$$

$$x(i)\in\{0,1\}, \qquad \forall i\in V, \qquad (4)$$

목적식 (1)은 선택된 노드의 가중치의 합을 최대화시키는 식이고, 제약식 (2)는 선택할 수 있는 노드의 개수에 대한 제약을 반영하는 식이다. 그리고 제약식 (3)은 노드 j가 선택되면 인접된 노드가 적어도 하나는 선택되어야 함을 의미하는 것이다. 제약식 (4)는 변수의 정수조건을 나타낸다.

3.2 상한의 도출

본 논문에서는 (IP)의 선형계획완화문제를 푸는 대신에 이 문제의 쌍대문제를 풀어서 상한을 도출하려고 한다. 왜냐하면 쌍대문제의 구조적 특성이 해를 구성하기가 수월하기 때문이다. (IP)의 선형계획완화문제는 정수조건 (4)가 다음의 제약식으로 변경된다.

$$x(i)\le 1,\quad i\in N \qquad (5)$$
$$x(i)\ge 0,\quad i\in N \qquad (6)$$

그러면 쌍대문제는 다음과 같이 구성된다.

$$(D)\quad \min\ \lambda k+\sum_{i\in V} u(i)$$

$$s.t.\ \lambda+w(i)-\sum_{j\in N(i)} w(j)+u(i)\ge d(i),\quad i\in V \quad (7)$$

$$\lambda,\ w(i),\ u(i)\ge 0,\qquad i\in V$$

여기서 $\lambda, w(i)$, $u(i)$는 제약식 (2), (3), (7)에 대응된 쌍대변수이다. 제약식 (7)의 실행가능조건을 표현하기 위하여 다음과 같은 기호를 정의하기로 하자.

$$s(i)=\lambda+w(i)-\sum_{j\in N(i)} w(j)+u(i)-d(i),$$
$$I^-=\{i\in N\mid s(i)<0\},$$
$$I^0=\{i\in N\mid s(i)=0\}.$$

우리의 해법은 (D)의 실행가능해를 구하여 원문제의 상한으로 이용하는 것이다. 이러한 쌍대기반의 상한을 구하는 기법은 정수계획문제에서 흔히 사용되는 방법이다[4]. 좋은 실행가능해를 구하기 위하여 다음과 같은 쌍대문제의 특성을 이용한다. 쌍대변수 $w(i)$는 목적함수에 나타나지 않는다. 쌍대변수 $u(i)$는 $s(i)$와 목적함수를 동일한 값만큼 증가시킨다. 쌍대변수 λ는 모든 $s(i)$를 증가시키지만 목적함수를 k배 증가시킨다. 우리의 목표는 목적함수의 값을 최소로 하면서 모든 $s(i)$가 비음의 값을 갖도록 하는 것이다. 따라서 우리의 전략은 가능한 $w(i)$를 이용하고, 여의치 않은 경우에 $u(i)$와 λ를 증가시키되, 음수의 값을 갖는 $s(i)$가 적어도 k개 이상인 경우에만 λ를 증가시킨다.

Algorithm UB

1. Initially set $w(i)=u(i)=0$, $i\in V$ and $\lambda=0$.
2. For each $i\in I^-$, if $s(j)>0$ for all $j\in N(i)$, then set $w(i)\leftarrow w(i)+\min\{|s(i)|, \min_{j\in N(i)} s(j)\}$ and update $s(j)$ for $j\in N(i)\cup\{i\}$ and I^-.
3. If $|I^-|>k$, set $\lambda\leftarrow\lambda+\min\{|s(i)|\mid i\in I^-\}$ and update $s(i)$ for $i\in V$ and I^-.
4. If any update in steps 2 or 3, go to step 2.
5. For each $i\in I^-$, set $u(i)\leftarrow|s(i)|$ and $s(i)\leftarrow 0$.

4. 상한의 도출

이 장에서는 정수계획모형의 실행가능해를 구하는 휴리스틱을 개발하려 한다. 우리의 해법은 노드를 추가해 가는 방식의 휴리스틱(add heuristic)이다. 실행가능한 해를 구하기 위해서는 선택된 노드 중에 고립된 노드가 없어야 한다. 이를 위해서 노드를 추가할 때에는 이미 선택된 노드와 인접하거나, 그렇지 못한 경우에는 인접하는 노드와 함께 추가될 수 있는 경우에만 추가되도록 한다. 노드를 추가하는 휴리스틱의 성과는 어떤 순서로 노드를 선택할 것인지에 따라서 크게 영향을 받는다. 우리는 두 가지 전략을 시도해 보았다. 하나는 상보여유조건을 반영하기 위하여 $s(i)$=0인 노드를 우선적으로 선택하고 동일한 조건인 경우에는 가중치가 큰 노드를 먼저 선택한다. 두 번째 실험해 본 전략은 $s(i)$에 상관없이 가중치가 큰 노드를 먼저 선택하는 것이다.

5. 계산실험 및 결과 분석

본 연구에서 제시된 고립된 노드가 없는 최대가중치 k-노드 부분그래프를 구하는 문제의 상한과 하한을 구하기 위한 절차는 C 언어를 통하여 프로그램으로 구현되었다. 계산실험에 사용된 대상문제는 실제 기지국을 표현한 네트워크와 가상으로 만들어진 네트워크를 대상으로 만들어졌다. 최대 선택 가능한 노드의 수, 즉 k는 전체 노드 수의 1/3 정도로 설정하였다.

작성된 코드는 1.7Ghz CPU를 가진 펜티엄급 PC에서 실행되어 졌다. <표 1>에 80문제에 대한 결과가 나타나 있다. 문제별로 주어진 그래프의 크기를 노드의 수로 표시하였다. 제시된 두 가지 하한을 구하는 방법을 비교하였으며 각 수치는 10개 문제의 평균치이다.

<참고문헌>

[1] CDMA2000 High Rate Broadcast- Multicast Packet Data Air Interface Specification, 3GPP2 TSG-C SWG2.5, Aug. 18 2003
[2] Broadcast-Multicast Service (BCMCS) Framework Draft Document, 3GPP2, Aug. 18. 2003
[3] 3GPP TS 22.146 Multimedia Broadcast/ Multicast Service: Stage1, 3GPP, Mar. 2003
[4] Nemhauser, G.L. and L.A. Wolsey, *Integer and Combinatorial Optimization,* Wiley, New York 1988.

<표 1 > 계산결과 (10문제의 평균치)

|V|	k	UB	LB		GAP
			전략1	전략2	
9	3	2.112	2.112	2.108	0.0006
18	6	4.223	4.181	4.182	0.0080
29	10	6.804	6.737	6.743	0.0084
36	12	8.462	8.36	8.368	0.0104
48	16	11.363	11.147	11.156	0.0172
69	23	16.242	16.204	16.055	0.0119
129	43	30.477	29.956	29.983	0.0157
228	76	53.742	52.773	52.875	0.0171

GAP=(UB-best LB)/best LB