Journal of The Korea Society of Computer and Information
Vol. 20 No. 12, pp. 101-106, December 2015
www.ksci.re.kr
http://dx.doi.org/10.9708/jksci.2015.20.12.101

# A Divide-and-Conquer Algorithm for Rigging Elections Problem

Sang-Un Lee*

## Abstract

This paper suggests heuristic algorithm with polynomial time complexity for rigging elections problem that can be obtain the optimal solution using linear programming. The proposed algorithm transforms the given problem into adjacency graph. Then, we divide vertices $V$ into two set $W$ and $D$. The set $W$ contains majority distinct and the set $D$ contains minority area. This algorithm applies divide-and-conquer method that the minority area $D$ is include into majority distinct $W$. While this algorithm using simple rule, that can be obtains the optimal solution equal to linear programing for experimental data. This paper shows polynomial time solution finding rule potential in rigging elections problem.

▸ Keywords : Majority, Minority, Adjacency graph, Divide-and-conquer, Rigging elections

## I. Introduction

선거에서 승리하는 방식은 우리나라와 같이 전체 선거구들 의 투표 결과를 합산하여 다득표자로 결정하는 방식과 미국과 같이 해당 선거구의 다득표자를 해당 선거구 승리자로 결정하는 방식으로 구분될 수 있다[1]. 하나의 선거구를 구성하기 위해서는 유권자수 충족 조건과 지역을 통합하는 조건을 병행하여 결정한다[2].

선거구를 획정하고, 선거를 진행하는 것이 일반적인 순서이다. 이 때, 자신에게 투표할 의향이 있는 지지자 수를 구역별로 알고 있다고 가정할 때, 자신에게 유리하도록 선거구를 획정하고자 하는 문제를 선거구 획정 문제 (rigging elections problem, REP)라 한다[2].

REP는 빈번히 제기되거나 이슈화될 문제는 아니나 헌법재판소의 "지역 대표성보다 인구 비례에 따른 표의 등가성이 더 중요하다는 취지의 판결"에 따라 2016년 치러질 대한민국의 제20대 국회의원 선거의 경우 선거구 재획정 문제가 엄청난 폭발력을 가진 이슈로 대두되었으며, 총선직전까지도 선거구 재획정이 될지는 의문시되고 있다.

만약, 지역구의 선거구 재 획정 계획이 있는 차기 선거에서 특정 후보가 차기 선거에서 승리하기 위해서는 타당성 있는 논리로 재 획정 의사결정에 영향을 미쳐야 자신에게 유리하도록 재 획정할 수 있을 것이다.

REP에 대해 현재까지도, 국내 실정에 적합한 연구는 실증 데이터가 없는 관계로 고려할 수 없다. 따라서 부득이 미국의 선거방식 관련 한정된 특정 데이터에 한정하여 문제를 정의하고, 해결책을 제시하고자 한다.

REP의 해를 얻기 위해서는 선형계획법 (linear programming, LP)[2]이나 집합 피복 문제 (set covering problem, SCP)/집합 분할 문제 (set partitioning problem, SPP)를 적용하고 있다 [3-8]. 그러나 SCP나 SPP는 Karp의 21개 NP-완전 (NP- complete) 문제들 중 하나로, 정확한 해를 다항시간으로 찾을 수 있는 알고리즘이 알려져 있지 않아, 부득이 유전자 알고리즘 (genetic algorithm, GA)과 같은 메타휴리스틱 기법을 적용하고 있다 [6,9]. 선형계획법은 최적화 대상과 제약조건에 대한 선형함수식들을 나열하고, 최적화를 시키는 수학적 방식으로 최적화 대상의 개수가 $m$개 인 경우 $O(m^4)$의 수행 복잡도를 갖는 관계로, 선형계획법 패키지를 활용하지 않고는 적용하기가 쉽지 않다.

이와 같이 검증 가능한 데이터 부족과 최근 관련연구의 부족에도 불구하고, 선거구 재 획정이 될 때 해당 지역구 후보자의 당선 여부를 결정할 수 있는 지대한 영향을 미치기 때문에 연

• First Author: Sang-Un Lee, Corresponding Author: Sang-Un Lee
*Sang-Un Lee (sulee@gwnu.ac.kr), Dept. of Multimedia Engineering, Gangneung-Wonju National University
• Received: 2015. 08. 31, Revised: 2015. 09. 15, Accepted: 2015. 09. 29.

구 주제로 선정하였다.

본 논문은 REP에 대해 LP보다 간단히 최적 해를 구할 수 있는 휴리스틱 알고리즘을 제안한다. 제안된 알고리즘은 자신에게 유리하도록 선거구를 획정하기 위해, 과반이상 지지 구역과 과반이하 지지 구역으로 양분하고, 과반이하 지지 구역을 과반이상 지지 구역으로 통합하는 분할정복 (divide-and-conquer) 기법을 적용하였다. 2장에서는 REP 개념과 Guéret 등[2]가 제시한 사례를 고찰해 본다. 3장에서는 REP에 대해 다항시간 복잡도로 최적 해를 구할 수 있는 휴리스틱 알고리즘을 제안한다. 4장에서는 제안된 알고리즘을 실제 데이터에 적용하여 알고리즘 적합성을 평가해 본다.

## II. Problem Description and Related Works

본 장에서는 다득표자 방식이 아닌 다득표 선거구를 해당 후보가 승리하는 방식을 적용한다. 대표적인 예로, 미국의 대통령 선거 방식은 그림 1과 같이 각 주별로 유권자 수에 비례하여 다른 점수를 부여한다. 따라서 50개 주의 총점은 538점이며, 특정 주에서 과반수 이상을 득표하면 해당 주의 점수를 획득하는 방식으로, 과반수 이상인 270점 이상 획득시 대통령으로 당선될 수 있다[10].

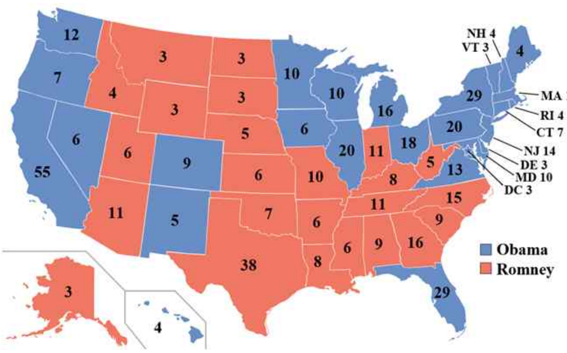



Fig. 1. Election method for president of U.S.A

이와 유사한 방식으로, Guéret 등[2]에서 인용된 그림 2의 경우를 고려하여 보자. 그림 2는 R. Tekris가 이끄는 연립여당과 맞서 싸우는 Duke 당의 S. Mevo에 대한 선호인원을 표현하고 있다. 이 문제는 현재 14개의 구역으로 분할되어 있으며, 단일 구역 선거구인 경우 50,000명 이상, 연합 구역 선거구인 경우 30,000 ~ 100,000명으로 선거구를 획정하고자 한다. 그림 2에서 단일 구역으로 선거구를 획정할 수 있는 50,000명 이상의 유권자를 보유한 구역은 2,4,10,12이며, 4,10은 승리할 수 있고, 2,12는 패배하여 5:5의 상황을 나타내고 있다. 따라서 나머지 10개 구역은 통합하여 선거구로 획정해야 한다. 여기서 다수의 구역이 하나의 선거구로 통합되기 위해서는 인접한 구역에 대해서만 통합되는 조건을 만족해야 한다.

Mevo는 Tekris와 경쟁하여 선거구를 획정하여 선거에서 승리하고자 한다. 이 경우 총유권자 수는 540,000명으로 최소 선거구 수는 6개가 된다. 즉, 6개 선거구로 획정시 Mevo는 과반 이상인 4개 이상의 선거구에서 과반 이상을 득표하여 승리하도록 해야 한다. 이 문제를 풀기 위해 Mevo는 수학자에게 도움을 요청하였다.

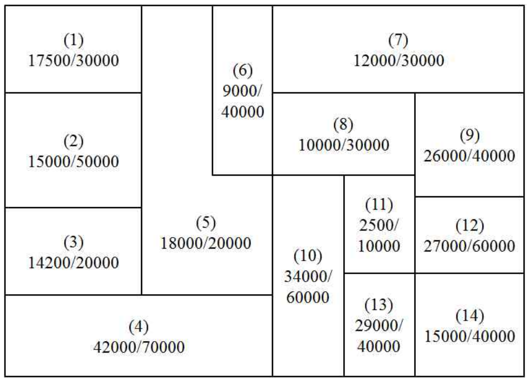
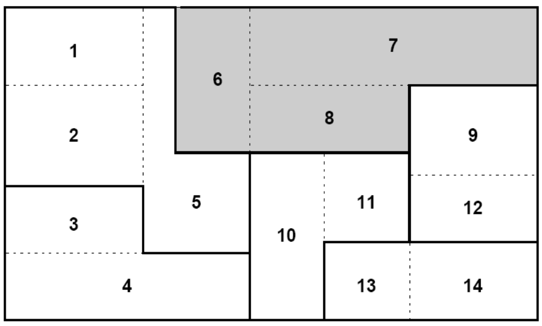



Fig. 2. Favorable votes/voters for districts of a capital

Guéret 등[2]은 그림 2의 문제에 대해, 선형계획법을 적용하여 그림 3의 결과를 얻었다. 즉, 6개 선거구로 획정한 결과 3 선거구를 제외한 나머지 5개 선거구에서 승리하여 5:1의 결과를 얻을 수 있었다. 이 문제는 집합 피복 문제/집합 분할 문제로 풀 수도 있다[3-8]. 그러나 SCP나 SPP 모두 NP-완전으로 정확한 해를 다항시간으로 풀 수 있는 알고리즘이 알려져 있지 않다[9].




| 선거구 | Distinct | Electors | Favorable votes | Rate (%) | 결과 |
|---|---|---|---|---|---|
| 1 | 1<br>2<br>5 | 30,000<br>50,000<br>20,000 | 17,500<br>15,000<br>18,000 | 58.33<br>30.00<br>90.00 | 승리 |
| | Total | **100,000** | 50,500 | **50.50** | |
| 2 | 3<br>4 | 20,000<br>70,000 | 14,200<br>42,000 | 71.00<br>60.00 | 승리 |
| | Total | **90,000** | 56,200 | **62.44** | |
| 3 | 6<br>7<br>8 | 40,000<br>30,000<br>30,000 | 9,000<br>12,000<br>10,000 | 22.50<br>40.00<br>33.33 | 패배 |
| | Total | **100,000** | 31,000 | **31.00** | |
| 4 | 9<br>12 | 40,000<br>60,000 | 26,000<br>27,000 | 65.00<br>45.00 | 승리 |
| | Total | **100,000** | 53,000 | **53.00** | |
| 5 | 10<br>11 | 60,000<br>10,000 | 34,000<br>2,500 | 56.67<br>25.00 | 승리 |
| | Total | **70,000** | 36,500 | **52.14** | |
| 6 | 13<br>14 | 40,000<br>40,000 | 29,000<br>15,000 | 72.50<br>37.50 | 승리 |
| | Total | **80,000** | 44,000 | **55.00** | |
| | | | | | **5:1** |

Fig. 3. Result of rigging elections for linear programming

3장에서는 분할정복 기법[11,12]을 적용하여 다항시간으로 해를 얻을 수 있는 휴리스틱 알고리즘을 제안한다.

## III. Divide-and-Conquer Algorithm

본 장에서는 다수의 구역을 하나의 선거구로 형성하기 위한 “인접 구역 통합 가능 조건”을 충족시키기 위해 구역을 정점 (vertex)으로, 간선 (edge)을 인접하는 경우로 표현한 인접 그래프 $G=(V,E)$ 를 적용한다. 다음으로, 인접 그래프 $G$에 대해 과반 이상 득표 (majority) 구역 $W$ 집합과 과반 이하 득표 (minority) 구역 $D$ 집합으로 양분한다. 마지막으로, 분할 그래프에 대해 과반 이하 득표 구역 $D$의 득표율 오름차순으로 인접한 과반 이상 득표 구역으로 병합 (정복)하여 선거구를 획정하는 방법을 제안한다. 제안된 알고리즘을 분할정복 알고리즘 (divide-and-conquer algorithm, DCA)이라 하며 다음과 같이 수행된다.

Step 1. 각 구역을 정점 $v\in V$로, 인접한 경우를 간선 $e\in E$으로 연결한 인접 그래프 $G=(V,E)$ 로 변환시킨다. 두 구역의 유권자수 합이 100,000 초과 간선은 삭제한다.

Step 2. 인접 그래프 $G$에 대해 과반 이상 득표 구역 $W$ (win) 집합과 과반 이하 득표 구역 $D$ (defeat) 집합으로 양분한다.

Step 3. $d\in D$의 득표율 오름차순으로 인접한 최대 득표율 $w\in W$로 1차로 통합 (병합)한다. 단, 통합 결과 과반 이상 득표율을 획득할 경우와 유권자수가 100,000명을 초과하지 않는 경우에 한해 수행되며, 그렇지 않으면 병합하지 않는다. 두 구역의 유권자수 합이 100,000 초과 간선은 삭제한다.

Step 4. $W$ 집합 내에서 단일 구역으로 선거구가 되지 못하는 50,000명 이하 유권자를 가진 구역을 유권자수가 적은 구역을 많은 구역으로, 최소 차수 구역부터 통합한다. 단, 이때는 통합된 선거구의 유권자수가 100,0 00명을 초과하면 안된다. 두 구역의 유권자수 합이 100,000 초과 간선은 삭제한다.

Step 5. 남은 $d\in D$의 단일 구역으로 선거구가 되지 못하는 구역이 존재하면 Step 3을 수행한다. 그렇지 않으면 알고리즘을 종료한다.

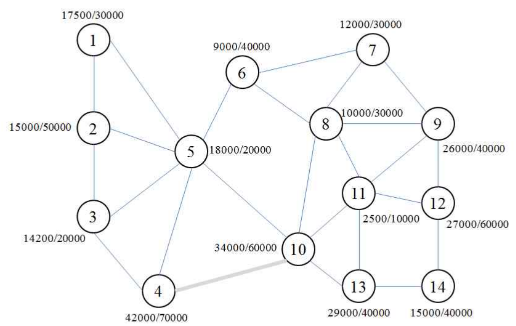
## IV. Applications and Evaluation

본 장에서는 3장에서 제안된 DCA의 5개 Step을 그림 2의 데이터에 순서대로 적용하여 보면서 DCA가 REP의 해를 찾아가는 규칙 (문제 풀이과정)이 될 수 있음을 보인다.

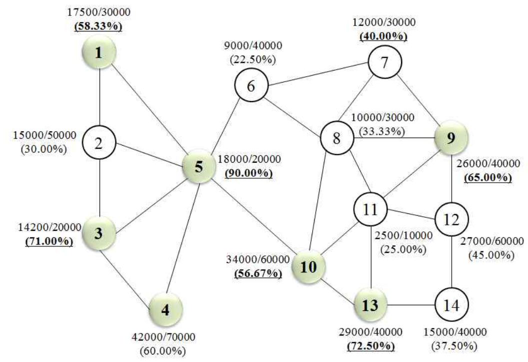
제안된 DCA를 그림 2의 데이터에 적용한 결과는 그림 4에 제시하였다. (a)는 Step 1을 수행한 결과로, 주어진 문제를 인접 그래프로 나타내고 여기서 4와 5 구역은 유권자수 합이 100,000명을 초과하여 통합될 수 없어 {4,5} 간선은 삭제되었다.

(b)는 Step 2를 수행한 결과로 $W$와 $D$ 집합으로 양분된 결과와 통합 결과를 보여주고 있다. 여기서, $W$= {1,3,4,5,9,10,13}, $D$= {2,6,7, 8,11,12,14}의 5:5로 양분되었음을 알 수 있다.

(c)는 Step 3을 수행하여 통합된 결과를 보여주고 있으며, $D$의 득표율 오름차순으로 최대 득표율 $W$에 병합되는 알고리즘이 수행된 결과이다. $D$의 득표율 오름차순은 12,7,14,8,2,11,6 순이다. 따라서 12→9 는 53,000/100,000=53.00%로 통합되며, 7은 통합 대상이 없으며, 14→13 은 44,000/80,000=55.00%로 통합된다. 8→10 은 44,000/90,000 = 48.89%로 통합되지 못하며, 2→5 역시 33,000/ 70,000 = 47.14%로 통합되지 못한다. 다음으로, 11→10 은 36,500 /70,000 = 52.14%로 통합되며, 마지막으로 6→5 는 27,000/60,000 = 45.00%로 통합되지 못한다. (c)의 결과에 대해, 통합시 100,000명 이상이 되는 간선이 삭제된 결과는 (d)에 제시되어 있다.

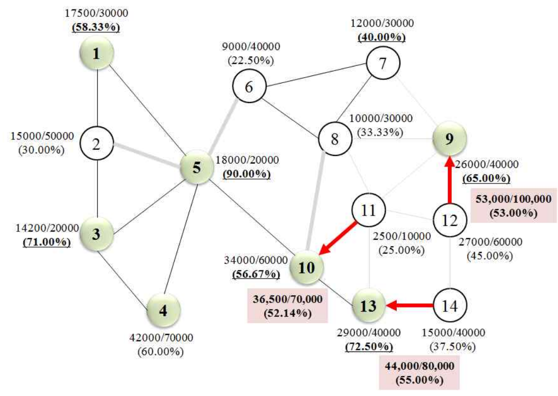


(a) Adjacent graph

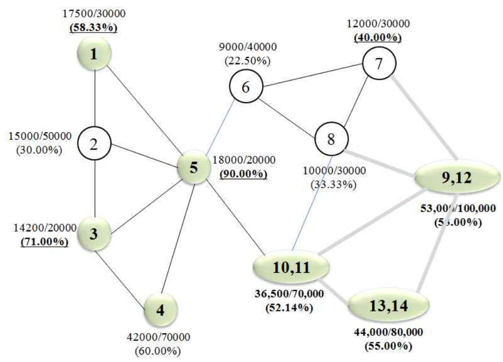


$W=\{1,3,4,5,9,10,13\}, D=\{2,6,7,8,11,12,14\}$

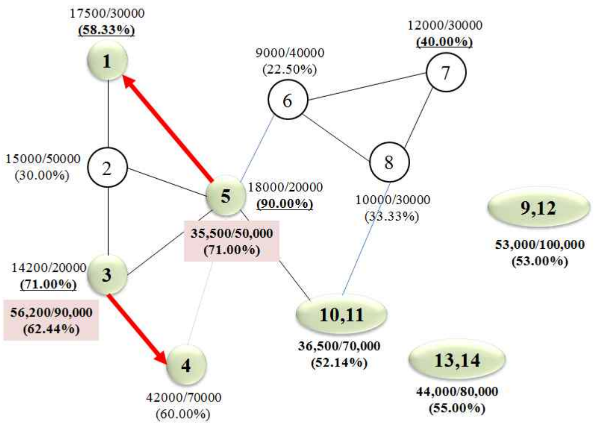
(b) Divide $V$ into $W$ and $D$



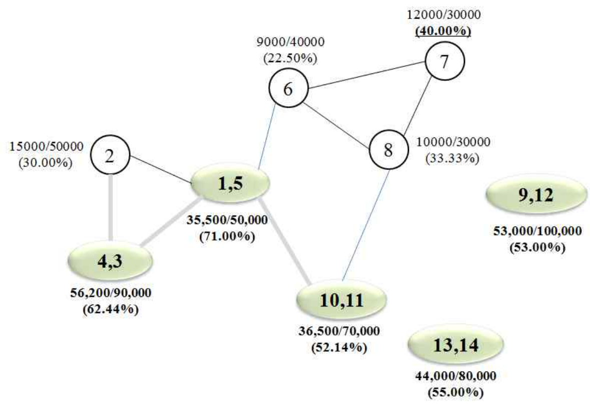
(c) 1st conquer for set $D$



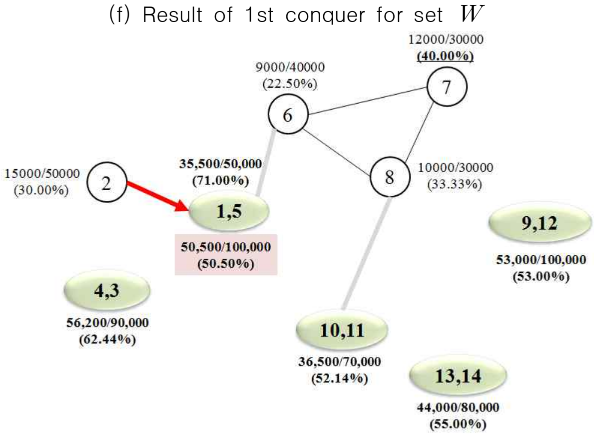

(d) Result of 1st conquer for set $D$



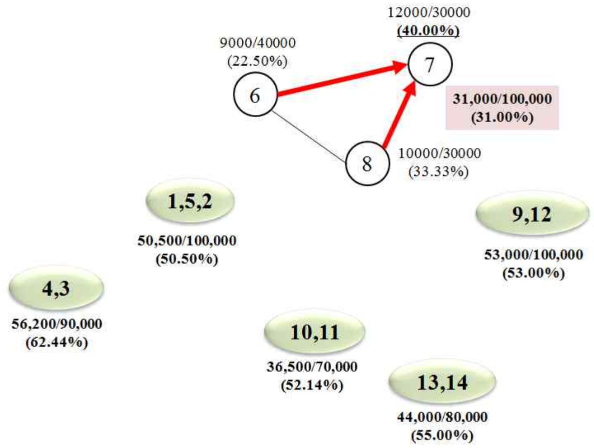

(e) 1st conquer for set $W$



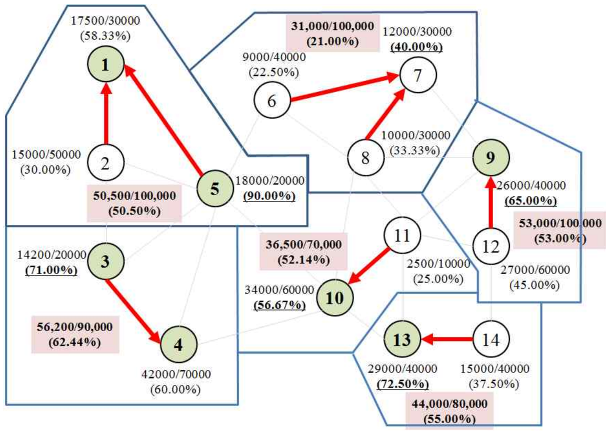

(f) Result of 1st conquer for set $W$




(g) 2nd conquer for set $D$




(h) 3rd conquer for set $D$




(i) Optimal solution

Fig. 4. Result of rigging elections for DCA

(e)는 Step 4를 수행하여 $W$ 집합 내에서 단일 구역으로 선거구가 되지 못하는 구역을 통합하여 선거구로 획정한 결과이다. 여기서, 단일 구역으로 선거구가 될 수 없는 50,000명 유권자를 갖는 구역은 1,3,5이며, 1은 5로, 5,3,4는 서로 연결되어 있어 최소 차수 구역 (정점)부터 통합하되 유권자 수가 적은 구역을 유권자가 많은 구역으로 통합하면 5→1, 3→4가 된다. 이 결과에 대해 선거구 최대 유권자수 100,000명을 초과하는 간선을 삭제한 결과는 (f)에 제시하였다.

(g)는 $D$를 $W$에 2차로 통합한 결과를 보여주고 있다. 여기서는 $W$에 인접한 간선의 개수인 차수가 최소인 구역 (정점) 우선으로 수행하면 2,6,8로, 2→{1,5}는 50,500/100,000=50.50%로 통합되며, 6→{1,5,2}는 100,000명을 초과하였으며, 8→{10,11}은 46,500/100,000=46.50%로 통합되지 않는다. 이 결과 남은 $D=\{6,7,8\}$은 $W$에 인접하고 있지 않고, 50,000명 이하로 독립적으로는 선거구를 구성할 수 없으므로, (h)와 같이 최대 득표율 구역인 7로 6,8이 통합된다.

결론적으로, 제안된 알고리즘은 (i)와 같이 6개 선거구로 획정하였으며, 5:1로 승리할 수 있음을 알 수 있다. 이 결과는 Guéret 등[2]의 LP 수행 결과와 동일함을 알 수 있다. 여기서는 LP의 수행과정은 설명이 불가한 관계로 생략한다. 왜냐하면, 선형계획법은 다수의 다변량 선형함수들의 합이 충족시킬 조건을 모두 만족시키도록 컴퓨터의 도움을 받아 시뮬레이션하는 관계로 제안된 알고

리즘과의 비교 설명이 불가하기 때문이다.

제안된 DCA와 LP의 성능을 비교한 결과는 표 1에 제시하였다.

Table 1. Compare with algorithm performance

| 알고리즘 | 구역 | 선거구 | 선거구별 결과 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | #1 {1,2,5} | #2 {3,4} | #3 {6,7,8} | #4 {9,12} | #5 {10,11} | #6 {13,14} | 결과 |
| LP $O(m^4)$ | 14 | 6 | 승리 | 승리 | 패배 | 승리 | 승리 | 승리 | 5:1 |
| DCA $O(mn$ | 14 | 6 | 승리 | 승리 | 패배 | 승리 | 승리 | 승리 | 5:1 |

제안된 DCA는 $O(mn)$ 의 다항시간 복잡도로 LP의 $O(m^4)$ 복잡도에 비해 보다 빠르면서도 동일한 결과를 얻는 장점을 갖고 있다. 따라서 수학분야의 전문적 지식을 보유하지 못하였거나 선형계획법 패키지를 활용하지 못하는 경우, 수학분야의 전문가 도움 없이도 제안된 알고리즘을 적용하면 선거에서 승리할 수 있는 선거구 획정 문제를 풀 수 있을 것이다.

## V. Conclusions and Future Research

본 논문은 단순히 선형계획법 패키지의 도움을 받아 해를 얻고자 한 선거구 획정 문제 (REP)에 대해 $O(mn)$ 의 다항시간 복잡도로 최적 해를 얻을 수 있는 알고리즘을 제안하였다. 즉, 본 논문에서는 REP의 해를 찾아갈 수 있는 다항시간 규칙 (휴리스틱 알고리즘)이 존재함을 보였다.

제안된 방법은 주어진 문제를 인접 그래프로 변환시키고, 정점들을 과반이상 득표 집합 $W$와 과반이하 득표 집합 $D$로 양분하여, $D$의 구역을 $W$ 구역으로 통합하여 선거구를 획정하는 분할정복 기법을 적용하였다.

이와 같이 제안된 알고리즘은 단순한 기법을 적용하였음에도 불구하고, 실험 데이터에 적용한 결과 LP와 동일한 최적 해를 얻을 수 있었다.

결론적으로, 제안된 알고리즘은 간단하고 쉽게 해를 구할 수 있는 관계로, 선거구 획정 문제에 직면한 경우 수학 전문가의 도움 없이 자체적으로 해결하는데 실제로 큰 도움을 줄 수 있을 것이다.

본 논문에서는 단지 하나의 사례에 대해서만 제안된 알고리즘과 기존의 선형계획법 결과를 비교하여 제안된 알고리즘의 적합성을 검증하고자 하였다. 추후 다양한 많은 데이터를 획득하여 제안된 알고리즘에 대해 보다 일반화된 적합성 검증을 수행할 예정이다.

## REFERENCE

[1] J. Bartholdi, C. Tovey, and M. Trick, "Voting Schemes for Which it can be Difficult to Tell Who Won the Election," Social Choice and Welfare, Vol. 6, No. 2, pp. 157-166, Apr. 1989.

[2] C. Guéret, X. Prins, and M. Sevaux, "Applications of Optimization with Xpress-MP: 15.3 Rigging Elections," Dash Optimization Ltd., pp. 234-227, Feb. 2005.

[3] J. E. Beasley, "An Algorithm for Set Covering Problems," European Journal of Operational Research, Vol. 31, No. 1, pp. 85-93, Jul. 1987.

[4] M. L. Fisher and P. Kedia, "Optimal Solutions of Set Covering/Partitioning Problems Using Dual Heuristics," Management Science, Vol. 36, No. 6, pp. 674-688, Jun. 1990.

[5] J. E. Beasley and K. Jφrnsten, "Enhancing an Algorithm for Set Covering Problem," European Journal of Operational Research, Vol. 58, No. 2, pp. 293-300, Apr. 1992.

[6] J. E. Beasley and P. C. Chu, "A Genetic Algorithm for the Set Covering Problem," European Journal of Operational Research, Vol. 94, No. 2, pp. 392-404, Oct. 1996.

[7] U. Aickelin, "An Indirect Genetic Algorithm for Set Covering Problems," Journal of the Operational Research Society, Vol. 53, No. 10, pp. 1118-1126, Oct. 2002.

[8] G. Lan, G. W. DePuy, and G. E. Whitehouse, "An effective and simple heuristic for the set covering problem," European Journal of Operational Research, Vol. 176, No. 3, pp. 1387-1403, Feb. 2007.

[9] R. E. Miller and J. W. Thatcher, Complexity of Computer Computations: Reducibility Among Combinatorial Problems," Plenum, pp. 85-103, 1972.

[10] Wikipedia, "United States presidential election," http://en.wikipedia.org/wiki/United_States_presidential_election, Wikipedia Foundation, Ltd., May. 2014.

[11] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, "Introduction to Algorithms," MIT Press, 2000.

[12] A. V. Levitin, "Introduction to the Design and Analysis of Algorithms," Addison Wesley, 2002.

## Authors

Sang Un Lee received the B. Sc. degree in avionics from the Korea Aerospace University in 1997. He received the M. Sc. and Ph. D. degrees in Computer Science from Gyeongsang National University, Korea, in 1997 and 2001, respectively.
He is currently Professor with the Department of Multimedia Science, Gangneung-Wonju National University, Korea. He is interested in software quality assurance and reliability modeling, software engineering, software project management, neural networks, and algorithm.