

Application of Computational Geometry on Multi-region Electoral Districting

Mawkae Hor*

‡School of Informatics
Kainan University
Taoyuan, Taiwan, ROC
mhor@mail.knu.edu.tw

*Corresponding author

Tzusui Hung

School of Industrial Management
Vanung University
Taoyuan, Taiwan, ROC
sharma@mail.vnu.edu.tw

Yingche Hung, Chengyuan Tang

Dept. of Information Management
Huafan University
Taipei, Taiwan, ROC
m4134@ms77.hinet.net
cytang@cc.hfu.edu.tw

Yifan Peng, Hungmin Hsu

Dept. of Computer Science
National Chengchi Univ.
Taipei, Taiwan, ROC
yfpeng@gate.sinica.edu.tw
tsbear1@gmail.com

Abstract—Different electoral districting affects the election results. The spirit of democracy is worthwhile that the election system must be fair and square and legitimately practices without the occurrence of that the political party and candidate taking advantages on the electoral districting inducing controversy over manipulating the election strategy. In this study, we developed a computational geometry based partitioning algorithm to solve the multi-region electoral districting problems. The algorithm uses the Geographic Information System as well as the dynamic programming techniques and solves the multi-region districting problems by recursively applying the two-partitioning algorithm. The problem regarding to the huge feasible solutions can be reduced substantially by introducing the concept of “indivisible regions”. The contiguity test and the compactness test can be done through the knowledge of computational geometry. The method is implemented on a local county to illustrate the entire mechanism and obtained feasible solutions for further evaluation and analysis. We also used the historical voting results to evaluate our districting results and compared with the results released by CEC (Central Election Commission). The analyses show that our methods can solve the multi-region electoral districting problems successfully and effectively.

Keywords—computational geometry; electoral districting; feasible solution; geographic information system

I. INTRODUCTION

Election is used to exercise the democracies over hundreds of years. A fair electoral system can accurately express the public opinions while different electoral districting could influence the election results. Traditionally, electoral districting was done manually that not only requires a lot of manpower but also material and time. The political parties tend to take advantages on biased districting and the manipulation of election strategy will worsen the objectivity and the fairness of elections.

A. Influences of Different Electoral Districting

Different electoral districting methods will inference the election results. The Gerrymandering phenomenon [1] perhaps is the most famous historical example of manipulating electoral districting occurred in 1812. We can see similar effects showing in Figure 1. In this figure, each circle and triangle represent one unit vote that favors the candidate from party A and party B, respectively. Both left and right subfigures have exactly the same voting behaviors but different districting methods. In the left subfigure, both parties receive a tight result. In the right subfigure, party A wins one seat while party B wins three seats.

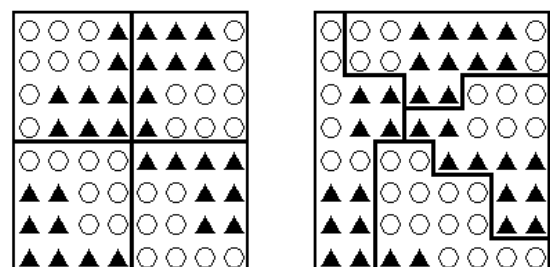


Figure 1. The influence of districting.

B. Principles of Electoral Districting

Various principles have been considered in electoral districting. According to the literature [2], the following three principles are the most important ones that every democratic country follows: (1) Population equality: the populations of every district should be similar; (2) Contiguity: all villages in a single district must stay connected; (3) Compactness: the shape of any district must be “good”. Additional principles that considered by various countries may include (4) Obey the existing administrative regions to avoid excessive partitioning;

- (5) Follow the natural boundaries such as mountains and rivers;
- (6) Respect the rights of the minorities.

Not all the principles can be satisfied at the same time. We will study only the most important three principles of electoral districting and meet the expectation of other principles through the “indivisible region” concept as much as possible.

C. Problems with the Electoral Districting

The CEC regulation requires using village as the smallest districting unit. A naïve algorithm in distributing balls into pockets will generally induce exponential complexity. For example, a naïve method of districting 100 villages into 3 districts will yield 3^{100} searching space. Exhausted search is generally not usable in this case. In addition, recursively applying the two-partitioning algorithm [3] may get huge amount of feasible solutions and can not be used to solve the multi-region districting problems directly.

II. RELATED RESEARCH

In the past, related research for electoral districting was traced to Harris’s rectangle method [4] in 1964 to meet and define the principles of districting methodology. Kaiser first presented a mathematical model to calculate the population equalization [5]. Hess’s linear programming method using algorithm involving restricted condition in optimizing the electoral districting [6]. Helbig improved on the linear program method by defining a formula to estimate the range of population deviation and then stepwise modify the districting result [7]. Leach proposed Divide and Conquer rules by recursive function to reduce the complexity of electoral districting [8].

Recently using algorithms in solving the electoral districting problems such as the Stimulated Annealing [9], the statistical physics [10], the surrounding method [11], the genetic algorithms [12], the centroidal Voronoi diagram and other electoral districting methodology for study is confirmed that different electoral districting could affect the election results [13]. Hsieh used the two-partitioning algorithm partitioning the simple two electoral districts successfully [14]; However, Hsu discovered that the two-partitioning algorithm to divide more than three electoral districts could induce the complexity problems on time & space [15].

III. MULTI-REGION ELECTORAL DISTRICTING

A. Formulation

In general, the districting problems can be formulated as distributing n villages into r districts. Let $V = \{v_i \mid i = 1, 2, \dots, n\}$ denote the set of n villages v_i . The population of village v_i is denoted as $P(v_i)$. Electoral districting is to divide V into r groups, $\{V_j \mid j = 1, 2, \dots, r\}$ so that all the village groups are disjoint and the population of each village group V_j , $P(V_j)$ are roughly the same. Note that the second condition requires that the population tolerance error, ε , measured in percentage of the average population \bar{P} , satisfies the CEC regulations.

This implies $\left| \frac{P(V_j) - \bar{P}}{\bar{P}} \right| \leq \varepsilon$, where \bar{P} is the average population of the electoral districts. Moreover, $P(V_j)$ and \bar{P} can be calculated as $P(V_j) = \sum_{v_i \in V_j} P(v_i)$ and $\bar{P} = \frac{1}{r} \left(\sum_{i=1}^n P(v_i) \right)$, respectively.

B. Multi-region Districting Solutions

We have developed the two-partitioning algorithm that divides a given county into two districts with equal population. Traditionally, we use $T(2, 1:1)$ to denote this partition, or, $T(2)$ in short. In general, we can use $T(2, i:j)$ to denote a two-partitioning algorithm that divide the given region into two districts with population ratio $i:j$.

Theoretically, multi-region electoral districting can be solved through repeatedly using the two-partitioning algorithm. For instance, three-partitioning can be obtained by first applying the two-partitioning process that partitions the original region into two districts with population ratio as 2:1. Then, follow by a second two-partitioning process that partitions the larger district into two sub-districts with equal population. In general, we have

$$T(n) = \begin{cases} T(2, i:j) + T(i) + T(j) & i \neq j, i+j = n \\ T(2, i:j) + 2T(i) & i = j, i+j = n \end{cases}$$

C. Problem with Multi-region Districting Solutions

The complexity of finding feasible solutions of n villages in the two-partitioning algorithm is roughly $O(n^3)$. Repeated execution of the two-partitioning algorithms in multi-region electoral districting has a lower bound at least in $O(n^4)$. The smallest n in our study is usually greater than 200, which easily generate over $O(10^{10})$ of feasible solutions. With this amount of feasible solutions, it is impractical for the further analysis.

IV. ARCHITECTURE AND METHOD

The system architecture of the electoral districting mechanism is shown in Figure 2. Data acquisition is mainly from the cartographic layers of Geographic Information System and the indivisible region has to be decided before passing the processed data for partitioning. The connectivity detection may be inserted after the looping and the compactness checking is done after all the regions districted to reduce the number of feasible solution. We exploit the Seventh Ballot information by CEC and sort out the substituted ballot votes from post electoral districting.

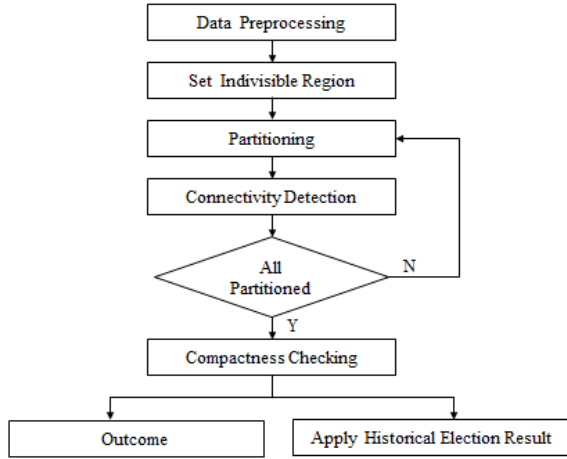


Figure 2. System Architecture.

A. Connectivity Detection

All villages in the same district must stay connected after the final partitioning. It can only control the difference in population of the districts within the tolerance range but not ensure that the district has no enclave condition. However, we can allow disconnected results provided the final results satisfy the contiguity principle during the intermediate process. The connectivity detection can be done using the depth-first search algorithm [16].

B. Compactness Checking

We use two criteria, the convex hull area ratio and the minimal circumcircle area ratio, to check the compactness of the districted regions. The convex hull area ratio, r_{CH} , defined as the ratio of the area of the district, A_D , to the area of the convex hull, A_{CH} , that contains the given district D . The minimal circumcircle area ratio, r_{MC} , defined as the ratio of the area of the district to the area of the minimal circumcircle, A_{MC} , that contains the given district D . That is,

$$r_{CH} = \frac{A_D}{A_{CH}} \text{ and } r_{MC} = \frac{A_D}{A_{MC}}.$$

It can be seen that both measurements are less than one and the closer to one the better the shape.

V. EXPERIMENTAL RESULTS AND DISCUSSION

Taoyuan County was chosen in illustrating our mechanism. Taoyuan County has 461 villages in 13 towns that have to be divided into six electoral districts. The city maps of Taoyuan are shown in Figure 3, the left subfigure shows the towns and the right subfigure shows the villages. According to census information, the population of Taoyuan City is approximately 1,853 thousands. The calculated \bar{P} is approximately 308 thousands. The regulations require $\varepsilon \leq 15\%$. The processing steps are described in the following subsections.

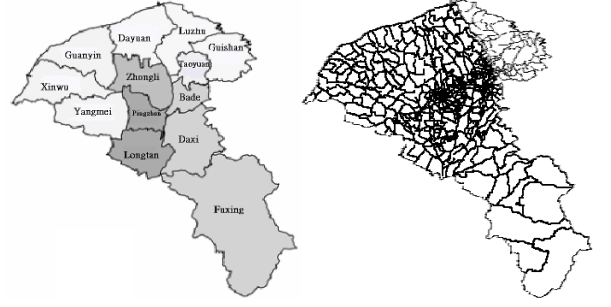


Figure 3. The map of Taoyuan County.

A. Partitions and Huge Number of Feasible Solutions

There are three different ways to apply the two-partitioning algorithms for the first partition step, $T(2,3:3)$, $T(2,4:2)$, or $T(2,5:1)$. $T(2,5:1)$ was chosen in this illustrations.

We first apply the partitioning algorithm that divides the Taoyuan City into two distinct districts using $T(2,5:1)$. The population tolerance ranges from 0.001% to 15% and the connected feasible solution is 5.6 million with $\varepsilon = 15\%$. Next we randomly chose any of the 4249 connected feasible solutions with $\varepsilon = 0.01\%$ and apply the partitioning algorithm that divides the district contains five units of populations into two distinct districts using $T(2,4:1)$. Again, we applied the partition $T(2,3:1)$ to the solution chosen from any of the 2411 connected feasible solutions with $\varepsilon = 0.01\%$. The final number of feasible solutions is summarized in Table 1. The counting number of feasible solutions after three consecutive partitions could reach 2931 billion. We know that just doing connectivity detection is not enough and the number of feasible solutions will increase substantially if no further restriction imposed on this mechanism. It is worthwhile to notice, regardless of how to set the error of population tolerance, the gained value of shape compactness ratio is very close and low value, that is to figured out the cause, it may be the population proportion of partitioning results is influenced by the effect of population density distribution

TABLE I. FEASIBLE SOLUTIONS AFTER $T(2,3:1)$, TAOYUAN COUNTY.

$\varepsilon(\%)$	Feasible Solutions		Ratio of Connectivity (%)
	Partitioned solutions	Connected solutions	
0.001	397	92	23.17
0.01	4,086	1,031	25.23
0.1	40,357	10,217	25.31
1	403,268	102,032	25.30
10	4,037,924	1,020,224	25.26
15	6,062,203	1,523,386	25.12

B. Indivisible Region

Indivisible region imposes additional constraints which can protect the rights of minorities and reduce the total number of the feasible solutions. Recall that Taoyuan County has 13 towns and the population of each town is listed in Table 2.

TABLE II. PAOPULATION OF EACH TOWN IN TAOYUAN COUNTY

Town Name	Population	Town Name	Population
Taoyuan	368,765	Bade	169,703
Zhongli	346,144	Longtan	110,222
Daxi	86,070	Pingzhon	198,353
Yangmei	136,178	Xinwu	49,593
Luzhu	118,581	Guanyin	56,300
Dayuan	79,404	Fuxing	11,011
Guishan	122,705		

The indivisible region can be set using the concept of onion peeling. Considering the geographic location and the boundary of every town, refer to Figure 4, we can peel and set the northern three towns, Dayuan, Luzhu and Kueishan as the first indivisible region. The population of this region is 320,690. Similarly, we set the western three towns, Guanyin, Xinwu and Yangmei as the second indivisible region, with population 242,071. Then, we set the southern three towns, Longtan, Dasi and Fusing as the third indivisible region, population 207,303. The population tolerance errors of these indivisible regions are 3.87%, -27.58% and -32.87%, respectively. Note that the population tolerance error in the second and the third indivisible region are over 15%.



Figure 4. The indivisible region (shaded area), Taoyuan city.

One can add villages to the regions in order to increase the populations to meet the requirement. Using this concept, the districting can be simplified from T(6) to T(3). Next, we apply the bricklayer method on Chungli town to make vertical partitioning, so that part of its villages will merge into the second indivisible region. We also make a horizontal partition to Pingjhen town so that part of its villages will merge into the third indivisible region, shown in Figure 4. The partitioning goes on to the remaining towns and villages to get the final three districts using T(2,2:1) and T(2,1:1).

C. Compactness Checking

We use the convex hull area ratio together with the minimal circumcircle area ratio to evaluate the compactness of the districted regions. Table 3 shows the number of feasible solutions after these checking of one particular experiment. The superscript “a” indicates the convex hull area ratio and the superscript “b” indicates the minimal circumcircle area ratio.

TABLE III. NUMBER OF FEASIBLE SOLUTIONS AFTER COMPACTNESS CHECKING, TAOYUAN COUNTY.

$\epsilon(\%)$	60% ^a	25% ^b	30% ^b	35% ^b	40% ^b	45% ^b
0.001	1	1	1	1	0	0
0.01	20	20	11	9	0	0
0.1	151	151	114	51	0	0
1	1,337	1,337	1,077	465	0	0
10	13,083	13,019	9,786	4,475	198	0
15	19,753	19,326	14,275	6,986	880	0

Figure 5 and Figure 6 show a particular solution among all the feasible solutions.

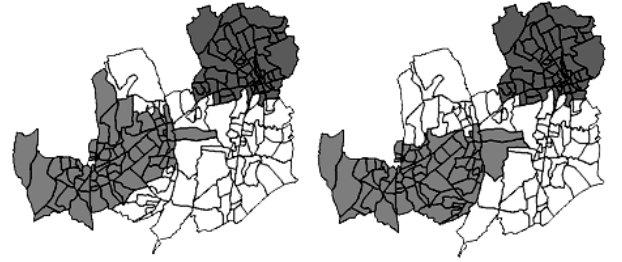


Figure 5. Example of districting results that pass the 60% convex hull area ratio test, Taoyuan county.

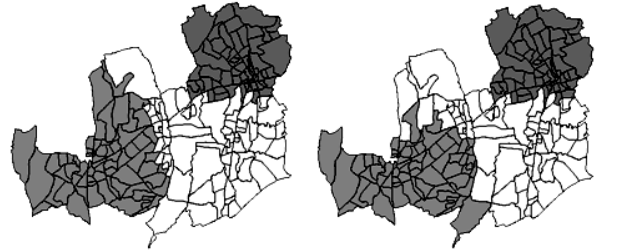


Figure 6. Example of districting results that pass the 60% convex hull area ratio test as well as the 30% minimum circumcircle area ratio test.

D. Historical Election Results Applying

Refer to Table 3, we randomly select 5 solutions from the 9 solutions that satisfy 0.01% population error, 60% convex hull area ratio, and 35% minimum circumscribed circle area ratio. These five methods, denoted as M1 to M5, are used to compare with the districts released by CEC. The results are shown in Table 4. In this table, district 1, district 2, and district 6 are the 1st, 2nd, and 3rd indivisible regions, respectively. The

number of historical votes stay the same in these districts regardless the districting methods. Also, district 1 is partitioned by using the bricklaying method and the historical votes stay unchanged too.

TABLE IV. COMPARISON OF CEC'S METHOD AND M1 THROUGH M5 BY APPLYING THE HISTORICAL VOTES

Districting Code	CEC	M 1	M 2	M 3	M 4	M 5
$\epsilon(\%)$	4.50	0.01	0.01	0.01	0.01	0.01
$r_{CH}(\%)$	70.21	67.48	63.99	73.31	60.50	60.62
$r_{MH}(\%)$	37.26	35.28	37.91	37.12	35.34	35.94
District 1 Blue-Vote	74026	72517	72517	72517	72517	72517
District 1 Green-Vote	44830	51012	51012	51012	51012	51012
District 2 Blue-Vote	72043	71490	71490	71490	71490	71490
District 2 Green-Vote	58503	58160	58160	58160	58160	58160
District 3 Blue-Vote	72134	82426	82426	82426	82426	82426
District 3 Green-Vote	44317	48476	48476	48476	48476	48476
District 4 Blue-Vote	82572	72921	73256	75727	72322	75995
District 4 Green-Vote	46610	34821	36522	34468	36145	36847
District 5 Blue-Vote	80898	85805	85470	82999	86404	82731
District 5 Green-Vote	41589	42584	40883	42937	41260	40558
District 6 Blue-Vote	85613	82127	82127	82127	82127	82127
District 6 Green-Vote	46423	47219	47219	47219	47219	47219
Elected-Blue-Seat	6	6	6	6	6	6
Elected-Green-Seat	0	0	0	0	0	0

One can compute the voting rate to the blue party and to the green party as well as the rate difference. This is summarized in Table 5. It can be seen that the gap is large enough to ensure the blue part to win in most of the districts.

TABLE V. TAOYUAN COUNTY THE SEVENTH ELECTION BALLOT

Electoral district	Total votes	Blue rate (%)	Green rate (%)	Difference (%)
1	118856	62.28	37.71	24.57
2	130546	55.18	44.81	10.37
3	122487	61.94	38.05	23.89
4	129182	63.91	36.08	27.83
5	122487	66.04	33.95	32.09
6	132036	64.84	35.15	29.69

VI. CONCLUSIONS

We proposed a systematic method that uses the techniques of computational geometry and GIS for multi-region electoral districting problems. This method cooperated with the two-partitioning algorithms, bricklaying method, onion peeling concept for indivisible regions, and the recursive techniques. We use the Taoyuan county as the example to illustrate our mechanism. Historical votes are used to analyze the districting outcomes and compared with the CEC's methods. The experiment results show that our method worked successfully.

ACKNOWLEDGMENT

The work in this paper is partially supported by the National Science Council, Taiwan, under the Grant No. NSC 101-2119-M-424-002 and NSC 101-2221-E-211-011.

REFERENCES

- [1] E. Gerry, "The Gerry-mander," Boston Gazette, 1812.
- [2] Central Election Commission, "Electoral Districting Rules for the Seventh Legislator Elections," Taipei: Central Election Commission, 2005. (in Chinese).
- [3] M. Hor, "Intelligent Electoral Districting Mechanism," International Conference of Machine Learning and Cybernetics (ICMLC), Qingdao, China, July 2010.
- [4] C. C. Harris Jr., "A scientific method of districting," Behavioral Science, Vol. 9, No. 3, pp. 219-225, 1964.
- [5] H. F. Kaiser, "A measure of the population quality of legislative apportionment," American Political Science Review, Vol. 62, No.1, pp. 208-215, 1968.
- [6] S. Hess, J. Weaver, H. Siegfeldt, J. Whelan, and P. Zitlau, "Nonpartisan political redistricting by computer," Operations Research Vol. 13, No. 6, pp. 998-1006, 1965.
- [7] R. E. Helbig, O.K. Patrick, and R. R. Robert, "Political redistricting by computer," Communications of the ACM, Vol. 15, pp. 735-741, 1972.
- [8] S. P. Leach, and K. Abraham, "Expert systems in government: a look at the redistricting problem," Washington University, Communications of the ACM, Vol. 15, pp.8, 1972.
- [9] C. I. Chou, and S. P. Li, "Statistical physics approach to political districting problem," Int. J. Mod. Phys. Vol. B 21, pp.3975, 2007.
- [10] C. I. Chou, and S. P. Li, "Taming the Gerrymander- statistical physics approach to political districting problem," Physics, Vol. A369, pp. 299, 2006.
- [11] M. Hor, C. Lee, K. Liu, et al., "Analysis of electoral redistricting," The 10th TAAI Conference on Artificial Intelligence and Applications, Kaohsiung, 2005. (in Chinese).
- [12] C. Lee, "Electoral redistricting in genetic algorithm," Master Thesis, Computer Science Department, National Chengchi University, 2006.
- [13] C. Wu, "Using Centroidal Voronoi Diagram in electoral redistricting,"
- [14] Master Thesis, Computer Science Department, National Chengchi University, 2009.
- [15] C. Hsieh, "Electoral redistricting using computational geometry," Master Thesis, Computer Science Department, National Chengchi University, 2008.
- [16] H. Hsu, "Multi-district electoral redistricting," Master Thesis, Computer Science Department, National Chengchi University, 2009.
- [17] Cormen T, Leiserson C, Rivest R et al., Introduction to Algorithms (2nd Edition). Cambridge: The MIT Press, 2001.