

# Improving Electoral Districting through Geographic Information Systems

Maw-kae Hor

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

Yi-Fan Peng

Center of GIS, RCHSS  
Academia Sinica  
Taipei, Taiwan, ROC  
d9607@cs.nccu.edu.tw

**Abstract**—In this paper, we present an approach that can solve the electoral districting problems by integrating the technologies of dynamics programming, computational geometry as well as the geographic information systems. A quantitative measurement to evaluate the shape of the districting results is also presented. We have applied these methods to Changhua County in Taiwan and obtained more than hundred thousands of feasible solutions.

**Keywords**—computational geometry; dynamic programming; electoral districting; geographic information system.

## I. INTRODUCTION

One of the most important issues in solving the electoral districting problem is to find a fair method in order to avoid the Gerrymandering phenomenon [1] as well as to meet the preset requirements.

In this paper, we present a mechanism that integrates the technologies from Geographic Information System (GIS), artificial intelligence, computational geometry, dynamic programming and other areas to solve the electoral redistricting problems systematically.

We used our method to district Changhua County of Taiwan to illustrate the entire mechanism. The experimental results show that our method can successfully obtain hundred thousands of solutions.

## II. RELATED STUDIES

Early research papers in this area can be found in 1960s. For example, Harris proposed rectangle method in solving the districting problems [2]. In 1965, Hess presented a linear programming method in optimizing the electoral districting [3]. Kaiser then focused on the population equalization by using the mathematical model to solve the districting problems [4]. Helbig improved Hess's method by defining a formula to estimate the range of population deviation and stepwise modify the districting result [5]. Leach presented divide and conquer rules applied in electoral districting [6].

Due to the change of election regulations in Taiwan, scholars paid more attentions to the elections systems and regulations as well as their influences. Hsieh and Pan discussed the mechanisms and standards on elections [7] [8]. Yu

presented a more complete discussion on the regulation changes and legal issues on it [9] [10]. Hsu gave some simulated results in predicting the election outcomes [11]. We also proposed various algorithms in attacking the electoral districting problem [12]-[14]. The goal in these researches is to provide a fair mechanism in electoral districting as well as to provide better electoral districting results.

## III. ELECTORAL DISTRICTING MECHANISMS

There are three most important electoral districting principles that most democratic countries follows:

- Population equality principle: the population in all districts should be more or less the same.
- Contiguity principle: the villages in the same district shall form a connected region.
- Compactness principle: the "shape" of any district shall be "good".

With these principles, we proposed the process of the electoral districting as shown in Figure 1. As indicated in the figure, data pre-processing has to be done before actual districting.

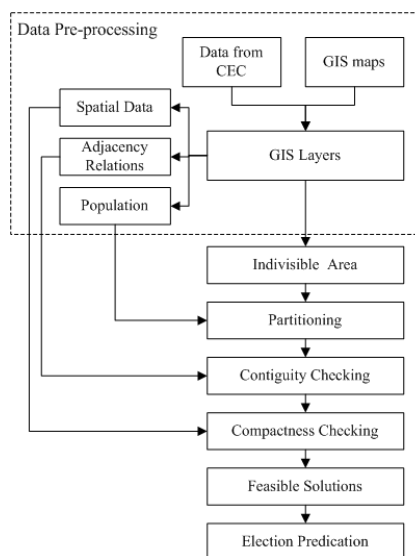


Fig. 1. The flow diagram of electoral districting

### A. Data Pre-processing

In the data pre-processing stage, we need to collect the raw data such as votes for every candidate from Central Election Commission (CEC). The election data is associated as the attribute information in GIS. The spatial information can be obtained through GIS. For example, the center coordinates, area, and perimeter information of each village  $v_i$ , will be needed during the partitioning stage. One can obtain the boundary coordinates, in sequential order, of each village,  $\{(x_j, y_j) | j=1, \dots, n_i\}$  from GIS map layer. Using the boundary coordinates, the village center,  $c_i$ , area,  $A_i$ , and perimeter,  $Pe_i$ , can be computed as

$$c_i = (c_{ix}, c_{iy}) = \frac{1}{n_i} \sum_{j=1}^{n_i} (x_j, y_j) \quad (1)$$

$$A_i = \frac{1}{2} \sum_{j=1}^{n_i} [(y_{j+1} + y_j)(x_{j+1} - x_j)] \quad (2)$$

$$Pe_i = \sum_{j=1}^{n_i} \|(x_{j+1}, y_{j+1}) - (x_j, y_j)\| \quad (3)$$

where  $n_i$  is the number of data points of  $v_i$  and  $j+1 = \text{mod}(j+1, n_i)$ .

We also need the adjacency relationships among the villages and among the towns for contiguity checking as well as repeated solution checking.

### B. Setting Indivisible Area

Besides the three principles mentioned at the beginning of this section, there are other factors one may want to consider during the actual districting process. For instance, it is suggested the actual districting must respect the natural boundaries as well as preserve the right of minorities. These considerations can be achieved through the concept of “indivisible area”. By setting the indivisible area, we could prevent unwanted separations during the districting processes.

### C. Repeated Solutions Checking

Refer to Figure 2 for a simple example of this problem.

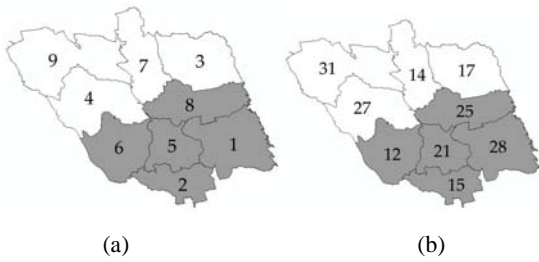


Fig. 2. Repeated districting results

In Figure 2(a), the entire region is districted in to two districts. The shaded district contains villages numbered 1, 2, 5, 6, and 8. The white district contains villages numbered 3, 4, 7 and 9. Similarly, in Figure 2(b), the shaded district contains villages numbered 12, 15, 21, 25, and 28. The white district contains villages numbered 14, 17, 27 and 31. These two

solutions represent the same districting result with different numbering scheme and a good districting mechanism should resolve this problem and keep only one of them in the final feasible solution set.

### D. Contiguity Checking

The contiguity principle requires that all villages in the same district must stay connected. We use, again, a simple example to illustrate this problem, as depicted in Figure 3. In this figure, the districting process districts the entire region into two districts, the shaded one and the white one. The shaded district contains villages numbered 2, 4, 7, and 9. The white district contains villages numbered 1, 3, 5, 6, and 8. It can be seen clearly that the shaded district is not connected where village 2 becomes a exclave. This yield an invalid districting results. However, further districting process may change the exclave unit into a connected result and yield a valid solution.

One can check the contiguity properties simply by using the depth-first search algorithm as well as the adjacency relations [15] and exclude the invalid solutions easily.



Fig. 3. The result with exclave problem

### E. Compactness Checking

We need to find a way when evaluating the shape of the district area. For instance, figure 4(a) is one of the district results. As you can see in this figure, the district D4 has strange shape. Villages of district D4 lied between district D1 and district D3. This strange shaped solutions can be avoided or removed by using the concept of “compactness evaluation”.

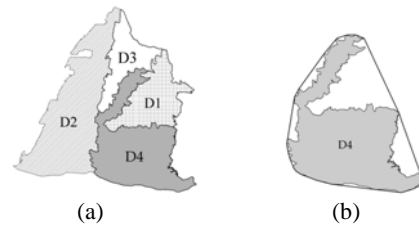


Fig. 4. The compactness checking problem

However, checking the compactness is not as easy as that of contiguity. Through the help of convex hull [16], for example, the surrounding line in figure 4(b), we define the area ratio and use it to measure the compactness of each district [17]. One can use  $A_D$  to denote the area of a district  $D$  and  $A_{CH_D}$  to denote the area of its convex hull. The area ratio of district  $D$ ,  $r_D$ , is defined as

$$r_D = \frac{A_D}{A_{CH_D}} \quad (4)$$

It can be seen that the best shape of a district occurs when the district is identical to its convex hull and, in this case, the area ratio equals to one.

Using this criterion, one can measure the compactness of every district. It is worth to note that we have to measure the compactness of all the districts generated by a particular districting method in order to conclude the compactness of this particular districting method.

#### IV. EXPERIMENTAL RESULTS

##### A. Experimental Data

We use Changhua County of Taiwan to illustrate the entire mechanism. Changhua County has twenty-six towns and five hundred and eighty-eight villages and will be districted into four districts as required by the election regulations. Choosing the villages as the smallest indivisible unit will generate innumerable feasible solutions using a naïve districting algorithm. For the convenience of illustration, we use town as the unit of districting just to reduce the number of feasible solutions. Fig. 5 shows the county map of Changhua with the townships indicated.



Fig. 5. The county map of Changhua county

TABLE I. THE TOWNSHIP POPULATIONS OF CHANGHUA COUNTY

Township	Pop.	Township	Pop.	Township	Pop.
Beidou	33,505	Huatan	46,336	Xianxi	17,010
Dacheng	20,063	Lukang	84,767	Xihu	56,356
Dacun	36,137	Pitou	32,316	Xiushui	38,431
Erlin	56,449	Puxin	35,082	Xizhou	32,584
Ershui	17,705	Puyuan	35,061	Yongjing	40,115
Fangyuan	37,980	Shengang	35,179	Yuanlin	127,042
Fenyuan	25,968	Shetou	45,591	Zhanghua	234,308
Fuxing	48,607	Tianwei	29,384	Zhutang	16,819
Hemei	88,245	Tianzhong	45,722		
<b>Average district population</b>				329,191	

The population of individual township is summarized in Table 1. The CEC regulations require that the population

tolerance error must not exceed fifteen percent. This gives a large margin during the actual districting process. We can choose a maximum population tolerance error and use it as a criterion to solve the districting problems. Note that the total population of Changhua county is about one million and thirty two thousands whereas the average district population is about thirty three thousands.

Fig. 6 shows the Changhua County districting maps announced by CEC. The information of individual district is summarized in Table 2. One can see the smallest area ratio is 65.21% and the largest population tolerance error is 8.21%, in absolute value.

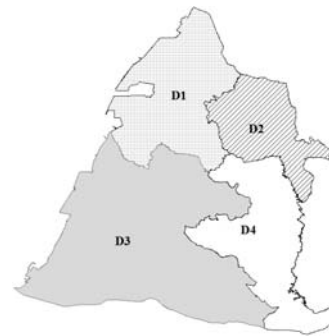


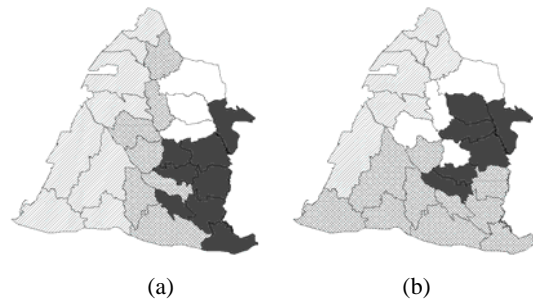
Fig. 6. The electoral district map of Changhua county by CEC

TABLE II. THE DATA OF INDIVIDUAL DISTRICTS (DISTRICTED BY CEC)

Electoral Districting	D1	D2	D3	D4
Area ( $km^2$ )	238.60	139.98	520.04	220.73
Convex Hull Area ( $km^2$ )	301.63	189.81	670.36	338.51
Area Ratio	79.1%	73.75%	77.58%	65.21%
Population tolerance error	5.15%	6.86%	-8.21%	-3.8%

##### B. Experimental Results

Using twenty-six townships of Changhua County, we successfully generated a large set of solutions by the mechanism we developed. Taking the 15% population tolerance error required by CEC, we obtained more than one hundred thousand feasible solutions. Figure 7 shown four randomly selected solutions.



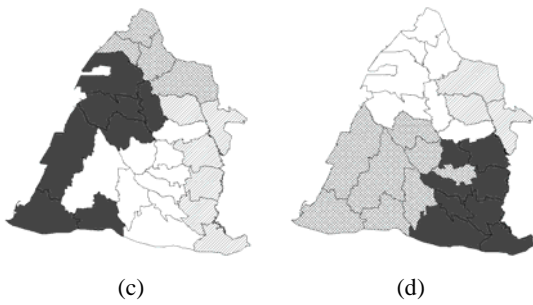


Fig. 7. Four randomly selected districting results

Though all the four results fit the requirements of CEC, one can see that the shapes of the districts are not good enough. One can use the area ratio to select solutions with better shapes. Fig. 8 shown four randomly selected solutions that satisfy the conditions  $-8\% < \varepsilon < 8\%$  and  $r_d > 70\%$ , for every individual districts. Using our performance evaluation criteria, these four solutions are better than the result announced by CEC.

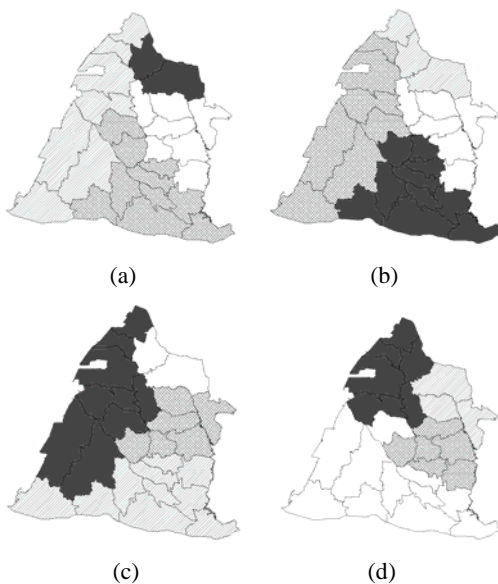


Fig. 8. Four randomly selected districting results with better area ratio

## V. CONCLUSIONS

In this paper, we presented a fair and nonpartisan mechanism for electoral districting. The methods we used integrated the technologies in GIS, computational geometry, artificial intelligence and dynamic programming. Principles of population equality, region contiguity, and region compactness are considered in our mechanism.

The experimental results indicate that our mechanism can generate a large set of feasible solutions for electoral districting. We can easily obtain results that are better than the CEC's districting results.

## ACKNOWLEDGMENT

The work in this paper is partially supported by the Ministry of Science and Technology, Taiwan, under the Grant No. MOST 103-2119-M-424-001.

## REFERENCES

- [1] W. Vickrey, "On the Prevention of Gerrymandering," *Polit. Sci. Quart.*, 67(1961):105-110.
- [2] C. Harris, "A Scientific Method of Districting," *Behavioral Science*, 9(3)(1964): 219-225.
- [3] S. Hess, J. Weaver, H. Siegfeldt and P. Zitlau, "Nonpartisan Political Redistricting by Computer," *Operations Research*, 13(6)(1965): 998-1006.
- [4] H. Kaiser, "A Measure of the Population Quality of Legislative Apportionment," *American Political Science Review*, 62(1)( 1968): 208-215.
- [5] R. Helbig, P. Orr, and R. Roediger, "Political Redistricting by Computer," *Communications of the ACM*, 15(8)(1972): 735-741.
- [6] S. P. Leach and A. Kandel "Expert systems in government: a look at the redistricting problem," *SIGAPP Applied Computing*. Rev. 1, 1 (January 1993), 2-9.
- [7] S. Hsieh, "Decision Making and Influence Factors of Electoral Districting of the Seventh Legislator Elections," *Annual Assembly and Conference*, TPSA, Taipei 2007. (in Chinese)
- [8] C. Pan, "Electoral System Reformation: from Electoral Districting Point of View," *Journal of Fu Hsing Kang*, 72: 237-268, 2001. (in Chinese).
- [9] C. Huang and C. Yu, "Evaluate the Change of Electoral Systems: Methodology Investigation," Wu-nan Book Inc., 2008. (in Chinese).
- [10] C. Yu, Y. Liu and Y. Wang, "On the Jurisprudence of Electoral Districting Problems after Amendments to the Constitution," *Research Report*, Ministry of Internal Affairs, 2006. (in Chinese).
- [11] J. Hsu, "Simulation of the Political Impact of Single District Two Votes Model," *Journal of New Century Think Tank*, 17(2002): 6-15. (in Chinese)
- [12] M. Hor, K. Chan, Y. Peng and H. Hsu, "Multi-region Electoral Districting using Computational Geometry," *International Conference of Machine Learning and Cybernetics (ICMLC 2013)*, Tianjin, China, July 2013.
- [13] M. Hor, T. Hung, Y. Hung, C. Tang, Y. Peng and H. Hsu, "Application of Computational Geometry on Multi-region Electoral Districting," *International Conference on the Modern Development of Humanities and Social Science*, Hong Kong, China, December 2013.
- [14] M. Hor, Y. Peng, M. Yang and C. Tang, "Districting Electoral Districts Intelligently," *International Conference on Computer Science and Network Security (CSNS 2014)*, Xian, China, April 2014.
- [15] Cormen T, Leiserson C, Rivest R *et al.*, 2001. *Introduction to Algorithms (2nd Edition)*. Cambridge: The MIT Press.
- [16] M. Shamos, "Computational Geometry," Ph.D. Thesis, Yale University, 1978.
- [17] M. Hor, "Intelligent Electoral Districting Mechanism," *International Conference of Machine Learning and Cybernetics (ICMLC 2010)*, Qingdao, China, July 2010.