INCORPORATING GENETIC ALGORITHM INTO SIMULATED ANNEALING BASED REDISTRICTING

Sim Kwan Hua

A thesis submitted
In full fulfillment of the requirements for the degree of
Master of Science
in Information Technology

Faculty of Information Technology
UNIVERSITI MALAYSIA SARAWAK
2002
BORANG PENGESAAHAN STATUS TESIS

JUDUL: Incorporating Genetic Algorithm Into Simulated Annealing Based Redistricting

SESI PENGAJIAN : 2000/2001

Saya SIM KWAN HUA
(HURUF BESAR)

mengaku membenarkan tesis * ini disimpan di Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dengan syarat-syarat kegunaan seperti berikut:

1. Tesis adalah hakmilik Universiti Malaysia Sarawak
2. Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dibenarkan membuat salinan untuk tujuan pangajian sahaja
3. Membuat pendigtit untuk membangunkan Pangkalan Data Kandungan Tempatan
4. Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi
5. ** sila tandakan (√)

☐ SULIT (mengandungi maklumat yang berdarah keselamatan atau kepentingan seperti termaktub di dalam AKTA RAHSIA RASMI 1972)

☐ TERHAD (mengandungi maklumat Terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

☑ TIDAK TERHAD

Disahkan oleh

(TANDATANGAN PENULIS) (TANDATANGAN PENYELIA)

Alamat Tetap:
No 158 Lot 77 Taman Wawasan
Lorong 15F Jalan Arang
93250 Kuching, Sarawak

Tanggal: 28/02/2003

Catatan: * Tesis dimaksudkan sebagai tesis bagi ijazah Doktor Falsafah, Sarjana dan Sarjana Muda
** Jika Tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh tesis ini perlu dikelaskan sebagai SULIT atau TERHAD
DECLARATION

No portion of the work refereed to in this report has been submitted in support of an application for another degree or qualification of this or any other university of institution of higher learning.

Sim Kwan Hua
00-02-0433
22 Feb 2003
Acknowledgements

I would like to express my deepest appreciation to my supervisor, Dr. Wang Yin Chai, for his encouraging words and for allowing me freedom to explore. His persistent guidance and patient has led to the success of this dissertation. For specialized knowledge and advice, I would like to convey my special thank to Dr. Stan Openshaw and his team members for their valuable conference proceedings. Thanks also to Dr. Kenneth Been for many fruitful email discussion on the implementation of Simulated Annealing. On a more personal note, I would like to thank my family for their unwavering encouragement and support throughout my research. Their love and prayers have me endured through all the difficult times. I am also feel very thankful to my great fellow colleagues for making my time doing this research both an enjoyable and rewarding experience.
# Table of Content

Acknowledgements ........................................................................................................... ii  
List of Tables .................................................................................................................... vii  
List of Figures .................................................................................................................... viii  
List of Abbreviations ........................................................................................................ xi
List of Abbreviations ......................................................................................................... xii  
Abstract ........................................................................................................................... xiii  
Abstrak ............................................................................................................................. xiii

## CHAPTER 1 INTRODUCTION

1.1 Background ................................................................................................................... 1

1.2 General Problem Statement ......................................................................................... 1

1.3 Objective ....................................................................................................................... 2

1.4 Scope of Study .............................................................................................................. 2

1.5 Research Methodology ............................................................................................... 3

1.5.1 Literature Review .................................................................................................... 3

1.5.2 Requirement Specification & Analysis .................................................................... 3

1.5.3 Designing Proposed Algorithm ............................................................................. 3

1.5.4 Implementation ...................................................................................................... 3

1.5.5 Testing and Evaluation ......................................................................................... 3

1.6 Significance of the research ....................................................................................... 4

1.7 Outline of the Dissertation ......................................................................................... 4

## CHAPTER 2 LITERATURE REVIEW

2.1 Introduction ................................................................................................................... 5

2.2 Redistricting ................................................................................................................. 5

2.3 Categories of Redistricting Factors ........................................................................... 6

2.3.1 Compactness ......................................................................................................... 7

2.3.2 Continuity .............................................................................................................. 9

2.4 Redistricting As A NP-hard Problem ........................................................................ 10

2.5 Existing Redistricting Algorithms ............................................................................. 12

2.5.1 Cake-cutting Algorithm ..................................................................................... 12

2.5.2 Greedy Algorithm ............................................................................................. 13

2.5.3 Steepest Ascent Algorithm ............................................................................ 13

2.5.4 Simulated Annealing ....................................................................................... 14

2.6 Genetic Algorithms .................................................................................................... 15

2.6.1 Fitness Landscape .............................................................................................. 15

2.6.2 Genetic Operators ............................................................................................ 16
2.6.2.1 Selection .................................................................16
2.6.2.2 Crossover ..................................................................16
2.6.2.3 Mutation ..................................................................18
2.6.3 Genetic Algorithms in Solving NP-hard Problem ..........18
2.7 Summary ........................................................................19

CHAPTER 3 SOLUTION DESIGN ........................................20
3.1 Introduction .................................................................20
3.2 Extended Analysis of Simulated Annealing .....................20
3.2.1 Analogy of Simulated Annealing .................................20
3.2.2 Outline of Simulated Annealing ..................................21
3.3 Specification of Simulated Annealing In Redistricting Process 22
3.3.1 The Basic AZP Algorithm ...........................................22
3.3.2 A Simulated Annealing AZP Algorithm .......................23
3.4 Algorithm Design of The Proposed Redistricting Algorithm ........................................................................25
3.4.1 Proposed Model of Incorporating Genetic Algorithm .........25
3.4.2 Incorporated Redistricting Algorithm ............................28
3.4.3 Reproduction Effect of Genetic Algorithm ......................29
3.4.4 Evolve Through a Set of Potential Candidate Solutions ........29
3.4.5 Annealing Schedule ...................................................31
3.4.5.1 Initial Temperature ..................................................31
3.4.5.2 Length of Iteration ..................................................32
3.4.5.3 Temperature Decrement Ratio .................................33
3.4.6 Selection Method ........................................................34
3.5 summary ........................................................................36

CHAPTER 4 IMPLEMENTATION ........................................37
4.1 Introduction .................................................................37
4.2 Implementation Environment ..........................................37
4.3 Test Data ....................................................................37
4.4 Simulation Development ................................................38
4.4.1 Initialization ............................................................38
4.4.1.1 Definition of Objective Function .............................39
4.4.1.1.1 Population Equality ...........................................39
4.4.1.1.2 Compactness ....................................................40
4.4.1.2 Preliminary Experiment ...........................................41
Incorporating Genetic Algorithm into Simulated Annealing Based Redistricting Algorithm

Abstract ........................................................................................................................................ 96

Keywords..................................................................................................................................... 96

1.0 Introduction .......................................................................................................................... 97

2.0 Redistricting ....................................................................................................................... 97

3.0 Simulated Annealing ......................................................................................................... 98

4.0 Simulated Annealing Based Redistricting Algorithm ................................................... 100

5.0 Genetic Algorithm .............................................................................................................. 102

6.0 Development of Proposed Redistricting Algorithm ....................................................... 103

7.0 Result and Discussion ......................................................................................................... 108

8.0 Conclusion .......................................................................................................................... 111

References ................................................................................................................................ 111
List of Tables

Table 2.1 Redistricting problems that are at least NP-hard (Altman, 1998) .................. 11

Table 3.1 Selection probability of individual solution ....................................................... 35

Table 4.1 Euclidean Measurement ................................................................................. 42
Table 4.2 Parameters Setting For Simulated Annealing Features ................................. 42
Table 4.3 Parameters Setting For Genetic Algorithm Features ..................................... 43

Table 5.1 Summary of redistricting simulation for modest districting plan .................. 54
Table 5.2 Summary of redistricting simulation for modest number of areas ................ 60
Table 5.3 Summary of this redistricting simulation for greater number of areas .......... 63
Table 5.4 Summary of this redistricting simulation for modest area size ...................... 66
Table 5.5 Summary of this redistricting simulation for greater area size ....................... 68
Table 5.6 Summary of this redistricting simulation for odd shape districting plan ........ 70
Table 5.7 Objective function value of output districting plan ...................................... 71
List of Figures

Figure 2.1 More compact district shape ................................................................. 7
Figure 2.2 Less compact district shape ................................................................. 8
Figure 2.3 Transformation of shapes. (Atlan, 1998) .............................................. 8
Figure 2.4 Geographical concern in redistricting ............................................... 10
Figure 2.5 Fitness landscape .............................................................................. 16

Figure 3.1 AZP redistricting algorithm .............................................................. 23
Figure 3.2 Simulated annealing AZP algorithm ................................................... 24
Figure 3.3 The concept of incorporating GA into SA .......................................... 26
Figure 3.4 Architecture Design of Proposed Redistricting Model .................... 27
Figure 3.5 Transition of area into adjacent district ............................................ 27
Figure 3.6 Incorporated redistricting algorithm ................................................ 28

Figure 4.1 Diagram 0 of the Data Flow Diagram ............................................... 38
Figure 4.2 Level 1 Data Flow Diagram for Process 1.0 ......................................... 39
Figure 4.3 Shape compactness of district .......................................................... 41
Figure 4.4 Level 1 Data Flow Diagram for Process 2.0 ......................................... 44
Figure 4.5 Command to select polygons ............................................................ 44
Figure 4.6 Command to select adjacent polygons ............................................. 45
Figure 4.7 Command to perform loop task on all the district ............................ 45
Figure 4.8 Level 1 Data Flow Diagram for Process 3.0 ......................................... 46
Figure 4.9 Subcommand in the STATISTIC command ....................................... 46
Figure 4.10 Syntax for calculating idea population size ...................................... 47
Figure 4.11 Command to obtain sum of population .......................................... 47
Figure 4.12 Command to dissolve polygon ....................................................... 47
Figure 4.13 Command to Join Info Table .......................................................... 47
Figure 4.14 Obtain standard deviation for population .......................................... 48
Figure 4.15 Syntax for calculating compactness value ...................................... 48
Figure 4.16 Syntax for calculation objective function value ............................... 48
Figure 4.17 Checking for contiguity of district .................................................. 49
Figure 4.18 Command to pass arguments .......................................................... 50
Figure 5.33 Output districting plan by enhanced algorithm

Figure 5.34 Simulation result on odd shape districting plan
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>SA</td>
<td>Simulated Annealing</td>
</tr>
<tr>
<td>T</td>
<td>Temperature Parameter</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographical Information System</td>
</tr>
<tr>
<td>PA</td>
<td>Perimeter/Area</td>
</tr>
<tr>
<td>FD</td>
<td>Fractal Dimension</td>
</tr>
<tr>
<td>EM</td>
<td>Euclidean Measure</td>
</tr>
<tr>
<td>POP</td>
<td>Population</td>
</tr>
<tr>
<td>AML</td>
<td>Arc Macro Language</td>
</tr>
</tbody>
</table>
Abstract

Redistricting is a process of drawing lines as a boundary, it plays an important role in the process of decision making on space and spatial allocation. In redistricting, the main problem to be solved is to find the districting plan that maximizes the value function involved. Thus, redistricting process can be characterized as a combinatorial optimization problem and considered to be computationally intractable or NP-hard problem, so getting trapped in local optimal and difficulty in obtaining the most optimal solution have become the great challenge of redistricting problem. Latest approach of solving redistricting problem using simulated annealing has shown a significant improvement with its ability to escape from local optimal. However, simulated annealing does not guarantee outstanding performance in escaping from local optimal though it posses that special capability. Therefore, this research will incorporate the advantages of genetic algorithm to overcome the deficiency of simulated annealing in the process of redistricting. The evolutionary feature of genetic algorithm that have potential to overcome some of the dilemmas of redistricting process are adapted to enhance the efficiency and performance of simulated annealing redistricting algorithm. At the same time, the strengths of simulated annealing will still being retained. Thus, the proposed redistricting algorithm has been enhanced by the reproduction effect of genetic algorithm in approaching more optimal districting plan with better objective function value during redistricting process. The implementation of simulation on the proposed model shows that genetic algorithm can be used to incorporate with simulated annealing in solving hard spatial optimization problem to achieve optimal result as a final output of districting plan. The outcome will contribute significantly to various applications and systems including political redistricting, business territory management and allocation planning. There are also several recommendations for future work as an extension of this research,
Abstrak

CHAPTER 1 INTRODUCTION

Redistricting can be characterized as a combinatorial optimization problem and it is a complicated process. The complexity will increase when it also involves other factors such as political, demographic, geographical, and environmental. A complete solution for redistricting seems to require an interdisciplinary effort. Many different methods have been used or suggested for finding optimal districts. However, redistricting process by itself is a NP-hard problem, current redistricting algorithms and conventional programming methods are not efficient, especially on implementing spatial redistricting (Altman, 1998). Thus, this research aims to incorporate the advantages of genetic algorithms to overcome the deficiency of existing redistricting algorithms, enhancement will be made on existing algorithms for redistricting process in order to increase the chances of obtaining near to optimal redistricting plan.

1.1 Background

Redistricting is a process of drawing lines as a boundary. It plays an important role in the process of decision making on space and spatial allocation. According to Agnew (1994), the spaces within which political, social, cultural, and economic processes are not only static backdrops or location referents for human events, but also the products of distinct territorial structures, identities, and ambitions, and they are deeply complicated by social and political change. Therefore, redistricting process plays an important role to redraw the lines on a district, based on the social and natural changes from time to time with the purpose of well spatial representation.

Thus, redistricting is extremely important in many different fields including ‘election boundary, school boundary, law enforcement, business territory management and even forest planning. Although each different redistricting process posses its own issues; such as gerrymanders for electoral boundary, some of the issues are consistent regardless of the redistricting domain; one of the biggest issue falls into this category is optimization in redistricting process.

GIS technology available today has offered a great chance for more consistent analysis of data and the ability to make more credible redistricting decisions in redistricting process (Hayes, 1996). Generally, optimization is one of the key factors in producing optimal and credible redistricting plan, but optimization could not be achieved completely with the improvement of technology. However, algorithms used for redistricting process can be improved in order to achieve the aim of optimization in redistricting.

1.2 General Problem Statement

In redistricting, the main problem to be solved is to find the districting plan that optimizes the value function involved. Formally, it is to find the optimal partition of the districting plan.

Redistricting process possess special difficulties because the size of the solution set can be enormous. Practically, it is impossible to solve the problem by a brute force search through all possible districting arrangements. As a result, redistricting process is considered to be computationally intractable or NP-hard problem. A problem will be said to be computationally intractable (also “computationally complex,” or “computationally hard”) if...
the optimal solution method for solving the problem cannot solve all instances in polynomial time.

Generally, current redistricting algorithms are still facing limited function and unable to satisfy spatial needs (Ngu, 1998). Besides, some of the current redistricting methods only approach to small size redistricting problems and others implement guessing procedures to get to near optimal solution. In solving bigger size of redistricting plan, the searching algorithms very often still not being able to approach near optimal solution and some of them will very likely to get trapped in local optimal.

Therefore, conventional programming methods for spatial redistricting algorithm not only has limitation on the implementation process but also facing inflexibility because of spatial complexity (Altman, 1998).

1.3 Objective

The research will focus on following objectives:

(I) Identify existing algorithms on redistricting and their weaknesses and problems in solving redistricting problem.

(II) Study the requirement and specification of redistricting algorithm in redistricting process.

(III) Investigate suitable features of genetic algorithm and their potential to enhance the redistricting algorithm that will increase the optimality of the obtained districting plan.

(IV) Design a redistricting algorithm by incorporating the evolutionary feature of genetic algorithms into existing simulated annealing based redistricting.

(V) Develop a redistricting simulation for the proposed algorithm.

1.4 Scope of Study

This research focuses on enhancing the current redistricting algorithm by incorporating the advantages of genetic algorithm. District boundary and compactness in redistricting process will also be the focus of the research in assessing the performance of the proposed algorithm.

Although there are numerous redistricting algorithms available currently, this research will only concentrate on simulated annealing based redistricting algorithm. The research will focus on the derivation of various components of the improved redistricting algorithm. The potential features of genetic algorithm are investigated in order to find the best way that they can be incorporated into simulated annealing based redistricting algorithm, and eventually increase the chances of obtaining district plan with better optimality.

Therefore, another main focus of this research will be on the genetic algorithm (GA) concept. Genetic algorithm uses process that analogous to natural evolutionary to generate solutions for an optimization problem. Thus, GA seems to have many characteristics that can be taken as an ideal solution on redistricting process to ensure better efficiency.

For redistricting simulation that will be implemented based on proposed model, latest Geographical Information System (GIS) technology will be used to assist the manipulation of spatial and attribute data during the implementation process. This is based on the
capability of current GIS technology in managing and handling spatial relationship besides powerful features and functions that allow spatial manipulation in the scope of study to be done.

1.5 Research Methodology

This research will be conducted in a systematic and orderly manner. There are five main phases in this research methodology, namely literature review, requirement specification, design, implementation and testing. This research methodology is to ensure that the research will achieve its objective systematically. The tasks of each stage in the methodology will be discussed as below:

1.5.1 Literature Review

Reviews will be done on current redistricting algorithms with extensive review on the simulated annealing based redistricting algorithm. At the same time, important redistricting criteria will also be identified. Genetic algorithms will also be studied and evaluated in order to understand its characteristic as well as its ability in solving optimization problems to discover the benefits of genetic algorithm on redistricting.

1.5.2 Requirement Specification & Analysis

Existing redistricting algorithm especially simulated annealing based redistricting algorithm will be investigated to identify the requirement and specifications that are needed to yield modification and improvement. Requirements of proposed algorithm will also be specified. Serious concern will also be given concerning the feasibility of the new technique to ensure that it will optimize the redistricting criteria.

1.5.3 Designing Proposed Algorithm

Based on the specification identified in previous stage, the most appropriate features of genetic algorithm will be incorporated into redistricting process to derive an enhanced redistricting algorithm, which will improve the capability of the redistricting algorithm in order to produce a more optimal district plan. Therefore, the detail specification and design of the proposed algorithm will be performed, and the proposed technique should fulfill the requirements and specifications of redistricting algorithm obtained from previous stage.

1.5.4 Implementation

A proposed redistricting algorithm will be implemented according to the specification requirements. A conceptual model on the proposed algorithm will be developed as a means to manipulate the proposed algorithm in redistricting process. In this stage, programming will be done in order to put the developed algorithm into implementation. A redistricting simulation of the proposed algorithm will also be developed so that it can be proven practically.

1.5.5 Testing and Evaluation

This is a simultaneous process with implementation. In this process, implemented solutions will be evaluated. The performance of the proposed algorithm will be tested and analyzed. Different sets of input data are used to evaluate the model in term of its effectiveness,
reliability, and performances. In the meantime, implementation alternatives, strength and weakness of the technique will also be identified and analyzed.

1.6 Significance of the research

No one can deny that the redistricting process plays an important role in making redistricting decision involving political, social and even business prospect. However, current redistricting algorithms and conversational programming methods seem to be not satisfactory. Thus, incorporating genetic algorithms is one of the ways to improve the optimality of the redistricting plan. This research should be benefitting all kinds of redistricting process and application. In short, this research should formulate a redistricting technique that makes redistricting tasks to be more effective in increasing the chances of finding near to optimal solution of redistricting plan.

1.7 Outline of the Dissertation

Every stage of the research will be documented. The outline gives a brief introduction of each chapter in this dissertation.

Chapter 2 describes the literature study of this research. It discusses the various elevation of redistricting algorithms, definition of redistricting, problems for the overall redistricting process, advantages and disadvantages of existing redistricting algorithms, possible approach to improve current methodology, and prospect of genetic algorithm in solving NP-hard problem will be reviewed.

Chapter 3 discusses the design of the proposed algorithm. It explains the development methodology. It contains an explanation on the solution and shows the general structure of the proposed algorithm. Various issues and factors are discussed throughout the designing process of the proposed redistricting algorithm.

Chapter 4 discusses the implementation of the proposed algorithm. It describes in details on the techniques of implementation. This section discusses the development framework of the implementation process. It describes all methods and techniques of implementation in details.

Chapter 5 discusses the results obtained from the implementation of the proposed algorithm with analysis from various aspects will also be carried. It explains the detail evaluation of the proposed algorithm by focusing on the incorporation of genetic algorithm into the algorithm of redistricting.

Chapter 6 discusses on the findings, accomplishments, limitations and future enhancement of the research. The achievement shows the contribution of the developed redistricting algorithm in improving redistricting process.
CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

This chapter aims to present the detailed literature review on redistricting from different perspectives related to this research. Redistricting algorithms are discussed thoroughly since it is one of the most crucial parts of redistricting process. This literature review tends to studies thoroughly existing redistricting algorithms, and identifies the requirements and specifications of the redistricting process. Among others, the review looks into, different kinds of algorithms and techniques in terms of their advantages and disadvantages, strengths, limitations, and the lacking in the redistricting process. These problems should stimulate the proposed domain in the following chapter as a solution to the problem. The existing algorithm will be incorporated with genetic algorithm, hence the potential features and suitability of genetic algorithm will also be reviewed. In short, the outcomes of this chapter are to conduct reviews and analysis on all aspects of redistricting, and later, to identify the specific problems of redistricting algorithm, and the potential of incorporating genetic algorithm for a better solution with better optimality district plan.

2.2 Redistricting

Redistricting process can be considered as an effort to explore the data for spatial relationships and representation. Although redistricting can be categorized into various kinds of categories based on the nature of data and their respective objective function of the redistricting process, this research tend to focus on redistricting that involves spatial data re-aggregation which is similar to Modifiable Areal Unit Problem (MAUP) defined by Openshaw and Taylor (1980), and viewed by Openshaw as one of the major redistricting problem.

One of the major task of redistricting process mentioned above starts with a set of data at one scale and then re-aggregate it again in order to create a new set of district plan suitable for a specific purpose. It involves the aggregation of spatial data for N areas into M districts, where M is less than N and is usually a small fraction of N in size. It is very important to note that all the districts should be internally connected and contiguous.

However, it is obvious that these activities will face problem due to the likely sensitivity of the result to the aggregation of the data being processed.

Expressed in mathematical notation, it looks like any other mathematical optimization problem trying to optimize F(Z), where F(Z) is a general function of Z, except Z is not a simple set of linear or non linear parameters, but defines an aggregation of N initial areas into M districts as output, and subject to constraints on Z to ensure the M district are internally connected, F(Z) is any function defined on data for the M districts in Z, and Z is the allocation of each of N areas to one of M districts such that each area is assigned to only one district.

In other words, there are implicit constraints on the function Z in such a way that each of the original N areas only can to be assigned to exactly one output district. Moreover, all the members from the same output districts have to be connected to form an individual district so that when the internal boundaries are dissolved, they will still be able form a single polygon.

For example, electoral redistricting can be considered as one of the important variant of redistricting, and Openshaw (1984) argued that it is a special case of a more general zone.
design problem that is equivalent to a special type of non-linear optimization problem in which the parameters to be optimized was the classification of N areas into M redistrict (N>M).

Unfortunately, there are millions of ways of aggregating N areas into M districts, and it is also important to note that different redistricting algorithms may not provide same result when being applied to same data. In fact, there is no convenient general purpose mathematical programming representation for this problem. It is some kind of discrete integer non-linear discontinuous programming problem that could well have many sub-optimal (Openshaw and Rao, 1995).

The only general-purpose solution strategy was likely to be some kind of combinatorial search heuristic. Thus, this optimization task might be categorized as a constrained non-linear integer optimization problem. It can only be solved via heuristic methods that may not find the global optimum result; indeed, there is no way of knowing whether there is a single global optimum result to be obtained (Openshaw and Rao, 1995).

The current algorithms were still not effective enough to draw compact district plan because of two main reasons. Firstly, some of the existing studies did not consider geographical aspects like compactness and continuity during the redistricting process. Secondly, most of the studies did not concern much on the decision making process in redistricting, especially those involving multiple criteria. The complexity of the geographical-related problem has made the redistricting process to generate odd or not-practical shape for district plan (Bong, 2000).

2.3 Categories of Redistricting Factors

Generally, factors to be considered in redistricting process can be categorized into two categories. The first category is application dependent criteria that may different between various kinds of redistricting applications. This criteria is very much depending on the goals or objectives of the redistricting process. For instance, redistricting process for election boundary will consider population equality as its main factor.

While, the second category is applications dependent criteria, this criterion is consistence over different redistricting application and domain. Two of the most concerned factors under this criterion are contiguity and compactness. These criteria refer to the geographical concern of the redistricting process. Contiguity receives much more attention then compactness does, while compactness is concerning the regularity of districts (Altman, 1988).

Both categories of criterions are related to one another. For example, although political boundary redistricting treats population equality as the main criteria in creating an election boundary, compactness and contiguity criteria of the district will not be ignored.

Thus, multiple criteria that include both of the application dependent and application independent factors like compactness should be included to consider an integrated compactness measurement. The consideration of application dependent criteria was compulsory whereby the consideration of application independent criteria was the unique factor to ensure the optimality of shape compactness (Bong, 2000).

In fact, shape based redistricting using the FMCDDM method called Fuzzy-AHP, to integrate the multiple criteria (including the compactness measurement), to produce an optimal compact district plan has been introduced (Bong, 2000). The Fuzzy AHP approach allowed
the integration and consideration of both application dependent and application independent factors in the redistricting process.

2.3.1 Compactness

Compactness refers to the geographical regularity of a district, and can be determined by considering the district’s appearance, and the area of dispersal (Altman, 1998). It can also be determined by an analysis of the function of the district. There are several mathematical ways to measure the elements of compactness as mentioned by Knight (1997) concerning mathematical analysis of shape, regularity of the spatial shape, and the regularity of the population distribution within the district. Thus, irregular geographical boundaries and significant land areas may be justified unequal district lines if such district lines follow a significant geographical feature or political subdivision boundary (Altman, 1997).

An analysis on the function of district may determine the compactness of the district. For election boundary redistricting, the district should be drawn to facilitate integrated communication between a representative and his constituents, and integrated opportunity for voters to know their representative and the other voters he represents (Altman, 1998).

Different applications of redistricting will have different definition for compactness. In political redistricting, compactness means population dispersion, but it refers to student distribution in school redistricting.

However, geographical compactness is one of the redistricting criteria that will be taken into consideration in most of redistricting process. For instance, population redistricting aims to obtain districting plan with most optimal distribution of population size among all the districts, but it will not ignore totally the geographically compactness of the districting plan.

Generally, geographical compactness is defined as how tightly a shape is "packed", and it is often used as a characteristic to describe shape (Shioke, 1998; Knight, 1997). As a first level of inquiry, a district's compactness may be determined by considering its appearance and the area of dispersal of the district (Altman, 1998). Figure 2. 1 shows an example of a more compact districting plan in term of district shape compared to districting plan in Figure 2.2.

![Figure 2.1 More compact district shape](image-url)
According to Altman (1998), most compactness measures claim to describe the shape of a district, it should require that any index of compactness give the same score to two districts that have the same shape. Blair and Bliss (1967) suggest that two objects should be said to have the same shape if they identical through translation, rotation and uniform scaling.

Further more, three types of shape distortion and manipulation have been recognized: dispersion, dissection and indentation (Blair and Bliss 1967; Flaherty and Crumplin 1992; Frolov 1974). As describe in Figure 2. 3, dispersion reflects the symmetry of a shape around its center, so a circle is evenly dispersed, whereas a ellipse is less evenly dispersed. Dissection reflects discontinuity in the distribution of points across the convex hull of a shape, and shapes with holes cut out of them are highly dissected. Lastly, Indentation reflects the smoothness of the perimeter of a shape, the shape shown in Figure 2. 3 with coastlines are examples of indented shape.

These three types of shape distortion are good proxies for geographical manipulation, then acceptable measures of geographic compactness should capture at least one, if not more, of these principles. Based on this believe, Altman (1998) has formalize these principles:
Let a shape $S = s_1, \ldots, s_i \{ \}$ be a finite, nonempty set of simple, continuous, closed, non-overlapping subsets of the plane where $\text{Area } s_i \cap s_j (\cdot) = \text{Perimeter } s_i \cap s_j (\cdot) = 0, \forall i \neq j$.

Let $P:S \to \mathbb{R}^+$ be the length of the perimeter of the shape, and let $A:S \to \mathbb{R}^+$ be the area of the shape.

Let a compactness measure $C$, be a function $C:S \to \mathbb{R}$.

Based on these definitions, Altman formally define what it means for a compact measure to capture shape:

1. *Scale independence*: if two shapes differ only in scale, then they should be equally compact.

2. *Rotation independence*: if $S_1, S_2$ are two shapes, which differ only in rotation around the origin, they should be equally compact.

3. *Translation independence*: if $S_1, S_2$ are two shapes which differ only in position, they should be equally compact.

4. *Minimal dispersion*: A compactness measure reflects the principle of dispersion if, for all shapes $S_1, S_2$, if $S_1$ and $S_2$ are of equal area, and the perimeter of the convex hull of $S_1$ is larger, $S_1$ is less compact. Compactness measures that claim to capture dispersion are usually based on the ratio of a shape’s perimeter to its area.

5. *Minimal dissection*: If $S_1$ and $S_2$ are any two shapes with identical convex hulls, and $S_1$ has a strictly smaller area, then $S_1$ should be judged as less compact.

6. *Minimal indentation*: If $S_1$ and $S_2$ have identical convex hulls and $S_1$ has a strictly larger perimeter/area ratio, $S_1$ should be judged as less compact.

A compactness measure must not violate any of the first three principles above. One can also use the convex hull to compare two shapes that have the same general outlines, so as to see which is relatively more dissected or indented.

### 2.3.2 Continuity

Continuity means the adjacency of the district and the district must not has detached part. The requirement is to ensure districts all inter-connected. It is one of the most important factors to be considered in the process of redistricting. A district is said to be contiguous if all parts of the district touch one other at more than one single point, so that the entire district is within a continuous boundary. Legal standards governing redistricting for governmental institution often require all of the territory in each district to be contiguous.
2.4 Redistricting As A NP-hard Problem

It cannot be denied that researchers in computer science and operational research have met an agreement that redistricting is computational complex or intractable, even many simpler redistricting sub-problems are likely to be intractable. Therefore, redistricting is categorized as a NP-hard problem which is non-deterministic polynomial (Burris, 1997).

A problem is said to be computationally intractable or computationally complex if optimal algorithm for solving the problem does not guarantee to solve all instances of that problem in polynomial time.

Cook (1971) defined the first NP-hard problem, which has now been shown to belong to a large set, consisting of hundreds of problems in many fields. Karp characterized that polynomial-time reducibility is the most important property of NP-hard problems (Karp, 1972). If a problem is NP and all other NP problems are polynomial-time reducible to it, the problem is NP-hard (Altman, 1995).

These are problems with non-polynomial time with exponential time complexity, most of the exact methods and algorithms fail to solve in reasonable time on such problems with typical size in practice. As a result, it has no chance but to give up the demand of getting exact solution or absolute optimum, and be satisfied with a near optimal solution.

In the theory of complexity, NP donates the set of all problems, solvable by non-deterministic polynomial time algorithm. Non-deterministic here means that no particular rules is followed to make the approximation.

Search for a proof of the intractability of NP-hard problems has became one of the most famous open problems in computer science for the past two decades. While no proof of intractability has been found, no polynomial algorithms have ever been found that solves any of these problems, and because of the breadth of the class of problems, it is widely believed that no such algorithms exist (Altman, 1998).

Altman (1995) has shown formally that most of the common redistricting sub-problems, such as finding an optimal set of compact districts, are NP-hard. Table 2. 1 summarizes