FUZZY MULTIPLE CRITERIA DECISION MAKING METHOD IN INTEGRATING COMPACTNESS MEASUREMENT FOR SHAPE-BASED REDISTRICTING

BONG CHIN WEI

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

FACULTY OF INFORMATION TECHNOLOGY
UNIVERSITI MALAYSIA SARAWAK
2001

DECLARATION

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

Bong Chin Wei
99-02-0333
19 APRIL 2000

ACKNOWLEDGEMENTS

I thank Dr. Wang Yin Chai, my supervisor, for his encouraging words and for allowing me freedom to explore. His persistent guidance and patient has led to the success of this dissertation. I would like to thank my family members who are indeed supportive throughout my research. Their love and prayers have me endured through all the difficult times. My beloved, Dr Lam Hong Yoong, provided a context of love, support, prayers and happiness, which has seen me through this degree. On the other hand, I am grateful to the Faculty of Information Technology folks for providing a friendly and relaxing environment. I am also very thankful to my great fellow colleagues, Mdm. Lau Bee Theng, Mr. Wang Yin Chung, Mr. Wong Pak Yew, Mr. Chai Ted Sing, Mr. Lee Beng Yong, Ms. Liew Lee Hung, Mr. Goh Koh Luong, Mr. Wong Ning Meng and others. Finally, I acknowledge with gratitude the financial support I have received from the *Pelan Tindakan Pembangunan Teknologi Perindustrian* (PTPTP) and Graduate Studies & Research Support Division, Universiti Malaysia Sarawak.

TABLE OF CONTENTS

Contents *page*

LIST OF TABLES

TABLES *PAGE*

LIST OF FIGURES

FIGURES *PAGE*

LIST OF ACRONYMS

Acronyms

FMCDM	Fuzzy Multiple Criteria Decision Making
MCDM	Multiple Criteria Decision Making
DP	Dynamic Programming
GIS	Geographical Information System
SDLC	System Development Life Cycle
OR	Operation Research
LW	Length v. Width
A	Area
PA	Perimeter/Area
OS	Other Shape Measures
FD	Fractal Dimension
EM	Euclidean Measure
POP	Population-based compactness measures
PL	Plan compactness measures
EDM	Euclidean Distance Mapping
ADD	Application Dependent Data Store
AID	Application Independent Data Store
SOR	Shape Optimal Rules
DPM	Data Preparation Module
SOP	Shape Optimal Rules
Fuzzy AHP	Fuzzy Analytical Hierarchy Process
DM	Decision maker (District Planner)
ICI	Integrated Compactness Indexing
COD	Combine Optimal District Module
TIN	Triangulated irregular network data model
AML	Arc Macro Language

ABSTRACT

Redistricting is not only used for political purpose like in election zone rearrangement, but also important for avoiding overcrowding in school, sales territory management in business and forest resources planning related applications and systems. Redistricting is a multiple criteria problem because it involves natural and synthetic factors at the same time. However, problems were found in redistricting techniques with the lack of consideration on the geographical factors like compactness and continuity that have caused odd and not-practical shape in district plans. Existing compactness measurement indices are vague and imprecise although there are more than thirty over compactness measurement methods used to assess geographical aspect of a district plan. These indices are used after redistricting process and there is no incorporation of compactness indexing in redistricting process. Therefore, the objective of this research is to generate a new compactness measurement indexing to improve the redistricting process especially in term of geographical aspects like district size, shape and boundary. The new indexing should be able to cope with fuzziness and multiple aspect of compactness measurement. In short, the research aims to obtain an optimal compact district plan.

In this research, Fuzzy Multiple Criteria Decision Making is integrated with multiple compactness measurements to produce the new compactness indexing, or namely Integrated Compactness Indexing (ICI) and it is then applied for a shape-based redistricting. Two compactness measurements, which include Fractal Dimension and Euclidean Measure with area-perimeter ratio, are selected to act as application independent criteria. An another criterion is also selected for an application dependent goal. Each criterion is indicated with their appropriate weighing vectors to indicate their importance for the redistricting process. Fuzzy Triangular Number is used because they are flexible, simple, and comprehensive with easy computation and efficiency. Therefore, the ICI representing the overall plan compactness is generated based on the synthesis of the concepts of fuzzy set theory, AHP, α-cuts concept and index of optimism of district planners to estimate the degree of satisfaction of the judgements on a district plan. Later, the compactness indexing is incorporated in shape-based redistricting process with the use of location/Allocation technique through Dynamic Programming.

A preliminary survey was conducted to define the shape optimal rules according to the selected criteria for the integrated compactness indexing before the implementation of the prototyping for the forestland blocking case study to verify the practicality of the proposed model. Several types of raw data are tested to analyze the effects to the methods and to evaluate the overall performance of the prototype in various conditions. The result shows that the definition of the shape optimal rules together with all variables like compactness indexing from each criteria, their weighing vectors, confidence level and attitude to risk needed to be used thoughtfully for the calculation of the ICI to ensure the optimality of the district plan (output). In conclusion, the simulation gives positive result because it shows that the integration technique manages to achieve optimum result in producing optimum compact districts. There are several recommendations for future work in the end of the thesis.

ABSTRAK

Pembahagian semula kawasan bukan sahaja digunakan untuk tujuan politik seperti dalam penyusunan semula bahagian pilihraya, tetapi juga penting untuk aplikasi atau sistem bagi mengelakkan kesesakan sekolah, perancangan pemasaran dalam sektor perniagaan, perancangan sumber hutan, dan sebagainya. Teknik pembahagian semula kawasan adalah masalah pelbagai kriteria kerana ia melibatkan faktor semulajadi dan faktor buatan manusia pada sesuatu masa. Walau bagaimanapun, masalah telah didapati dalam teknik pembahagian semula kawasan dalam kekurangan pertimbangan faktor seperti kepadatan dan kesinambungan yang menyebabkan kawasan berbentuk ganjil dan tidak bermakna dalam sesebuah pelan kawasan. Indeks kepadatan kini adalah kabur dan tidak tepat walaupun terdapat lebih daripada tiga puluh jenis cara ukuran kepadatan digunakan untuk menyemak aspek geografi bagi sesebuah pelan kawasan. Oleh itu, objektif bagi kajian ini adalah untuk menghasilkan satu indeks ukuran kepadatan yang baru untuk memperbaiki teknik tersebut terutamanya dari aspek geografi seperti saiz, bentuk dan kerumitan sempadan sesuatu kawasan. Indeks yang baru sepatutnya berkebolehan mengambilkira kekaburan dan pelbagai aspek dalam ukuran kepadatan. Pendak kata, penyelidikan ini bertujuan untuk mendapat satu pelan kawasan yang optima padat.

Dalam penyelidikan ini, Fuzzy Multiple Criteria Decision Making telah digabungkan dengan beberapa ukuran kepadatan yang lain untuk menghasilkan satu indeks kepadatan (ICI) yang baru dan ia kemudiannya digunakan untuk pembahagian semula kawasan berasaskan bentuk. Dua ukuran kepadatan seperti Fractal Dimension dan ukuran Euclidean dengan nisbah luas kepada perimeter dipilih sebagai kriteria yang tidak bergantung kepada jenis aplikasi. Satu lagi kriteria yang dipilih untuk tujuan yang bergantung kepada aplikasi. Setiap kriteria akan diberi vektor pemberat yang sesuai untuk menandakan paras kepentingannya terhadap proses pembahagian semula kawasan. Fuzzy Triangular Number telah digunakan kerana ia lebih mudah diubahsuai, senang, dan lengkap dengan pengiraan yang lebih mudah. Dengan itu, indeks ukuran kepadatan yang akan mewakili kepadatan pelan yang menyeluruh akan diperolehi dengan pemaduan konsep yang berasaskan Fuzzy Set Theory, AHP, indeks optimistik daripada perancang kawasan bagi menganggar pertimbangan mereka mengenai paras puas hati. Selepas itu, indeks kepadatan ini akan digabungkan dalam proses pembahagian semula kawasan dengan location/Allocation teknik melalui Pengaturcaraan Dinamik.

Satu siasatan permulaan dijalankan untuk menghasilkan peraturan bentuk optima berdasarkan kriteria yang dipilih sebelum penghasilan satu prototaip untuk kes pembahagian tanah hutan bagi menunjukkan kebolehgunaan model yang dicadangkan itu. Kemudian, beberapa jenis maklumat kasar diuji ke atas teknik tersebut untuk menganalisa kesan ke atasnya dan juga menunjukkan kebolehan prototaip ini dalam pada keadaan yang berlainan. Keputusan menunjukkan bahawa definisi bagi peraturan bentuk optima bersama dengan semua pembolehubah seperti indeks kepadatan yang dihasilkan daripada setiap kriteria, paras kepentingan yang diisytiharkan terhadap setiap kriteria, paras kepercayaan diri dan sikap terhadap risiko yang diperlukan perlu digunakan dengan cekap untuk pengiraan ICI supaya pelan kawasan yang dihasilkan mencapai paras optimum. Kesimpulannya, keputusan-keputusan simulasi yang diperolehi memberi keputusan yang positif kerana teknik pergabungan yang dicadangkan itu dapat menghasilkan kawasan yang optimum padat dalam proses pembahagian semula kawasan. Di penghujung tesis ini, beberapa cadangan dikemukakan untuk kerja-kerja masa akan datang.

CHAPTER 1: OVERVIEW

Redistricting is a complicated process because it reacts to political, geographical, demographic, cultural, or government factors and decisions. However, current redistricting algorithms face problems in implementation of spatial redistricting and inefficiency of conventional programming methods (Altman, 1998). It is a complex system and requires an interdisciplinary approach to its solution. This research aims to study current redistricting algorithms and the assessment methods for district plan called compactness measurement in order to develop a redistricting model. This research proposes the integration of Fuzzy Multiple Criteria Decision Making in compactness measurement index and use it for spatial-based redistricting. Therefore, this research focuses on the enhancement of the redistricting algorithm on the performance and efficiency for generating optimal compact district plan. The outcome of this research will give significant benefits to various applications and systems such as school districting and allocation planning, business districting and political district identification and management.

1.1 Background

Redistricting or districting draws lines for boundary with aims to foster community and it answers the tough question: where to draw the lines? (Jones, 1998). These terms are always used in election boundary application for drawing lines for the political district. However, redistricting process is not only relevant to election system, but it is needed in school, law enforcement and even forests planning related systems. Redistricting is tremendously important that reacts on how a government takes control on the space of a particular region. The spaces within which political, social, cultural, and economic processes unfold are not simply static backdrops or locations referents for human events, but the products of distinct territorial structures, identities, and ambitions, and they are deeply implicated in social and political change (Agnew, 1994). Therefore, redistricting process is important to redraw the lines of a district based on the social and natural changes from time to time.

Redistricting election boundary is the process of redrawing the physical boundaries of jurisdictions within states and localities (Ngu, 1998). This process assures the electoral districts are divided equally among residents based on population data from the census count (Terral, 1996). It assures that the districts are compact and contiguous. In Iowa, state law and court opinions require that political districts have nearly equal populations and be as contiguous and compact as possible. Liittschwager (1973) explains that political districts not only allow state and local governments to provide citizens with fire departments, schools, and law enforcement, but determine who can vote for whom in local, state, and federal elections. Since districts have such great impact, the task of creating them in redistricting process is important.

On the other hand, redistricting in school boundary is a consequential process. This process follows from the concerns for children because it seeks to arrange a school system's resources to ensure that every child can access to a high-quality education. Crowded schools and low-test scores are problems in educational institution in some countries today (Jones, 1998). Therefore, redistricting is now being considered in order to reassign pupils to the new school to reduce enrollments in schools that currently operating over its capacity. Redistricting will allow more children to go to schools closest to their houses and it will help in cutting transportation costs (Newsom, 1997). Besides, Brian & William (2000) presents evidences of the effect on house values of a school redistricting in Shaker Heights, Ohio in 1987. In short, neighborhood schools are disrupted, bus transportation is introduced, and school racial composition changes. In short, redistricting is not only important for political district, but also for school districting and allocation planning.

Besides, a police department is creating new boundaries that can avoid beat configurations with potential accessibility problems like identifying a major river as a physical barrier. Law enforcement agencies in Philadelphia experienced a development on a redistricting plan of crime analysis applications that visualize both spatial and temporal urban conditions with deployment of advance technology (Robert & Kevin, 1997).

In forestry resources management, redistricting is applied to forestland classification. Normally, forestland is classified by slope into different district called compartment and the compartment is broken down into smaller district called block using known features like rivers, roads, or license boundary as block boundary (Galaune & Mozgeris, 1998). The other criteria involve soils type and typical volume distribution.

Moreover, business district uses redistricting process to have sales territory management. It re-aligns sales personnel to more efficient sales areas. By then, customer service is improved and revenue per coverage area will be increased when sales territories are optimized. Improved sales territories will result in timesaving and allow sales representatives to follow up more new business opportunities.

1.2 Geographical Concerns In Redistricting

The important factors related to the geometric shape of a district include continuity and compactness. For instance, judicial court in United States holds that districts should not be oddly shaped and that all pieces of land in a district should be inter-connected (Handley & Maley, 1997; Knight, 1997). Although the shape of these districts is not the basis for several of court's decision in most of the country, the compact districts are important in creating the district boundaries. Therefore, the research will discuss the meaning and definition of these criteria in the following section in order to clarify the important terms in this research.

1.2.1 Compactness

Compactness is defined as how tightly a shape is "packed", and it is often used as a characteristic to describe shape (Shiode, 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). Then, Knight (1997) describes this includes a mathematical analysis of the shape, regularity of the spatial shape, and how regularly the population is distributed within the district. Therefore, irregular geographical boundaries and significant land areas may justify unequal district lines if such district lines follow a significant geographical feature or political subdivision boundary (Altman, 1997a).

Compactness may be determined by an analysis of the function 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). Thereafter, different applications will treat compactness with various meanings. For examples, compactness refers to population dispersion in political redistricting, but it refers to student distribution in school redistricting. However, the basic definition for compactness that will be used in this research is referring to geographical compactness which regarding to the shape of the district.

1.2.2 Continuity

Secondly, continuity is another popular geographic principle in redistricting. Many state constitutions in the United States list continuity as a requirement for legislative districts (Handley & Maley, 1997). The requirement is to ensure all districts to be inter-connected. Altman, (1997b) divides districts into three categories in order of divergence from real world continuity: practically contiguous, questionably contiguous, and non-contiguous. As in figure below, practical contiguous are all districts those are formally contiguous or depart from continuity because of islands off the coast of the district. Questionably contiguous is second category districts that are otherwise contiguous but that contained islands that are not directly off the coast of the district, districts that are non-contiguous but could be connected by straight bridges, and districts that are connected only by "points" (Altman, 1998). Districts which are non-contigous are non compact.

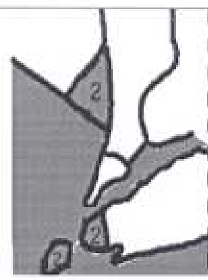
		
Non-contiguous	**Questionable contiguous** (it is connected only at a single point)	**Questionably contiguous** (the island portions of the district are not joined to the nearest mainland district)

Figure 1.1: Three different districts with their continuity category.

1.3 General Problem Statement

Thousand of district plans are generated from the redistricting process without knowing the best alternative. Therefore, there is a lack of knowledge-based system or effective decision making tool for automated redistricting process. Some of computer algorithms for redistricting can produce numerous computer-generated plans. However, the problem lies on the selection of the best plan. For example, *location/allocation* and *simulated annealing* approaches (refer Section 2.3) are able to generate 20,000 plans within one day (Carmen, 2000), but there is no study on selection of the best plan. On the other hands, generally employed large aggregate units may reduce the complexity of the problem. Another serious impediment to some of the redistricting approaches is the degree to which the results hinge on the start points. In short, there is a lack of research study on decision making process in redistricting.

Secondly, the natural geography of earth that comprises of forests, rivers, hills and mountains creates a situation that makes the process of creating the boundaries difficult (Ngu, 1998). Consequently, existing redistricting algorithms and studies do not attempt to solve spatial related problem in redistricting process. However, geographical criteria are important because redistricting without considering the natural boundaries will certainly create not-practical district such as a major river or a huge mountain in a district. Not-practical district will then create problem in transportation in election boundary redistricting and in school redistricting problem. In other word, there is necessary to consider geographical aspect in redistricting process and it causes redistricting problem becomes multiple criteria problem.

There are many ways proposed to measure compactness and they are commonly named Compactness Measurement. Altman (1998) even mentioned that there are more than thirty over measurement methods to evaluate district shape compactness but all of the measurement methods are used to assess the compactness of districts after the redistricting process. In other words, the measurement methods are used for checking of the redistricting result and there is no incorporation of compactness measurement into redistricting techniques to obtain optimal compact district.

1.4 Research Objectives

The overall objective of the study is to develop a fuzzy redistricting algorithm by integrating Fuzzy Multiple Criteria Decision Making in compactness measurement with the help of GIS technology to enhance the process of shape-based redistricting. The specific objectives of the research will be as following:

- Study current algorithms and methods on redistricting especially in terms of geographical aspect
- Identify spatial (shape) redistricting requirements and specification
- Integrate multiple criteria decision making technique in a compactness measurement indexing
- Apply the integrated compactness measurement indexing for a shape-based redistricting
- Develop a redistricting system (prototype) based on the proposed model

1.5 Scope of Study

This research focuses on enhancing the shape-based redistricting process by using an improved compactness measurement indexing which is integrated with Fuzzy Multiple Criteria Decision Making method due to the consideration of multiple criteria related to spatial and non-spatial during decision making process in shape-based redistricting problem. Then, the FMCDM method incorporates with fuzzy logic concept in order to simplify the decision making process in compactness measurement indexing and redistricting problem because fuzzy logic reflects how people think.

Factors like geographical features will certainly influence the result of redistricting especially the compactness of the district plan. With the multiple criteria mentioned, this research considers shape regularity or compactness and other geographical features like district size, shape, boundary and others, which are important in shape-based redistricting. Then, this study focuses on the efficiency, performances, and effectiveness of the model based on the results for the case study which will be selected among redistricting application like school, election, law enforcement or others.

This study will use current advance technology of Geographical Information System (GIS) in a prototype system according to the proposed model. This technology verifies the complexity of the spatial data and its attribute data management because GIS provides a lot of useful management in spatial relationship, powerful and ready-to-use spatial manipulation capability or functions related to shape-based redistricting and compactness measurement. Thus, the research can easily manipulate the geographical features in the scope of this study.

1.6 Research Methodology

The methodology in this research consists of the common phases in the System Development Life Cycle (SDLC) like analysis, design and implementation of a prototype system. Besides, the analysis and design methodology is by top-down-bottom-up approach as shown in Figure 1.2. There are three main stages in the methodology and the list of activities is defined in order to organize the research planning and development. The activities represented in box show the detailed flow and series for the methodology used in this research. The series proceeds with review and feedback in order to validate the research study from phases to phases.

The objectives of the stage one in the research framework are to acquire the understanding on the knowledge of redistricting process and techniques, the existing compactness measurement method, and the usage of GIS technology to the process. The research studies the backgrounds of the existing redistricting methods and the compactness measurement methods in order to grasp the knowledge on how the improved model will work. Therefore, this stage consists of the analysis of the present study, research and advanced development through searching or literature review of information from a variety of related materials such as Internet, reference books, journals and articles. By understanding of the current studies, the research able to identify strengths and weaknesses of the current algorithms and their common considered factors especially in term of geographical aspect. Besides, the research has a prediction of the future, showing that there is a demand to improve decision making process and spatial context in redistricting. Thus, this study identifies the objectives, the feasibility, and the requirements of the proposed model.

Based on the stage one above, the requirements and scopes of the research are distinguished. With the activities in stage one, the most appropriate technique and model are selected in order to design a suitable model for more effective redistricting process based on the requirements in previous stage. This stage is important to design a guideline for the model will be. Therefore, this research conducts a preliminary specification and design before the detail specification and design. Then, this stage will integrate the idea of the existing knowledge and to find the best solution needed for the proposed conceptual design and later for the proposed prototype system.

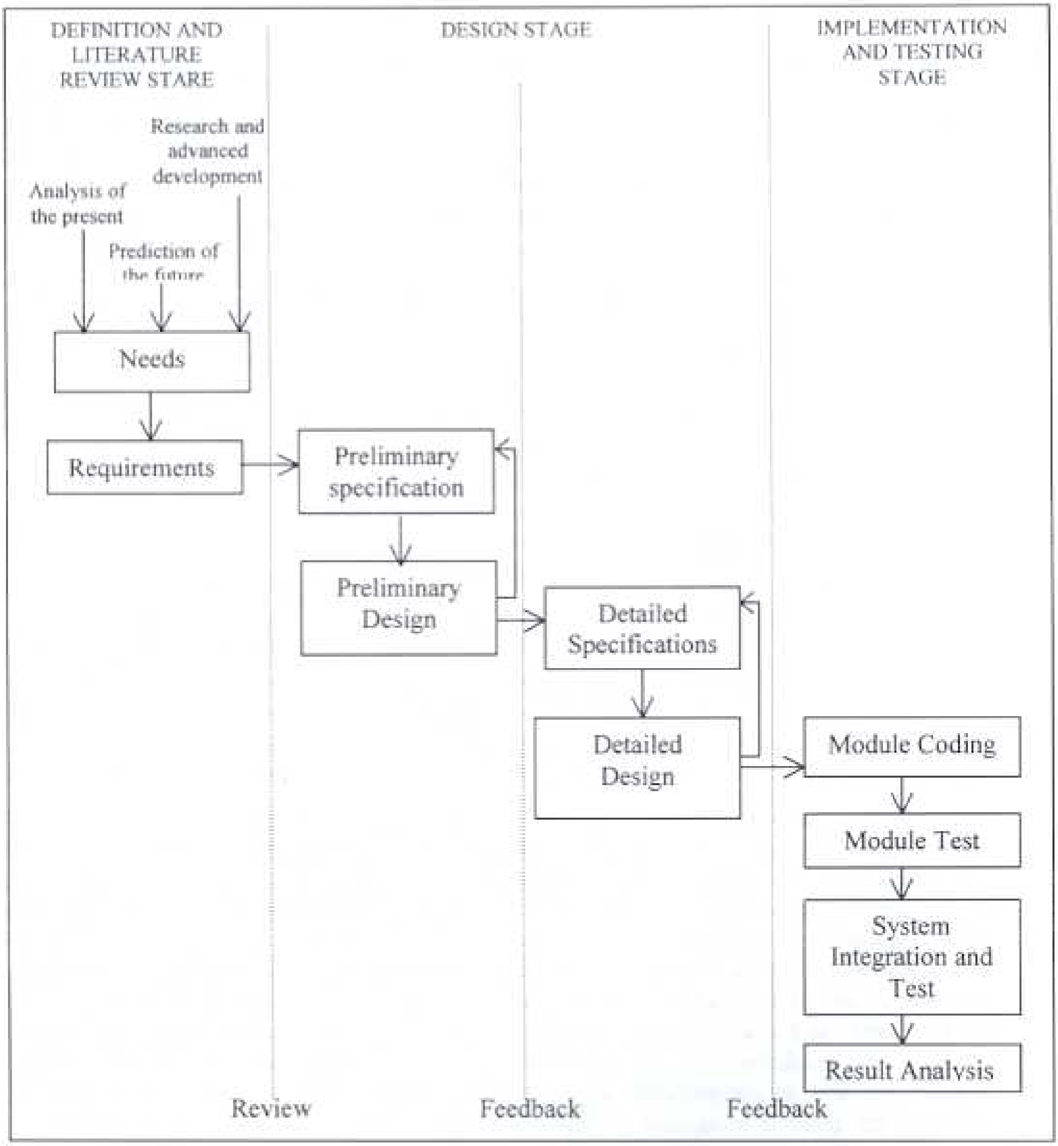


Figure 1.2: Methodology in this research

Then, the research will develop a prototype system based on the conceptual design, selected software, and programming language. In the stage three, some tools will be used in order to make the prototyping easier. First, a suitable case study and the necessary software tools are determined. Secondly, the research implements the design from the previous stage through coding by module. Then, the research uses modular testing so that this research can improve the prototype. Later, the research integrates all the modules and conduct overall testing to the prototype. Finally, the prototype system will be evaluated to correct possible error and to assess the actual performance of the prototype, which represents the proposed model. Different sets of input data are used to evaluate the model in term of its effectiveness, reliability, and performances to generate an optimal compact district plan.

1.7 Expected Output

At first, the expected outcome will be a thorough discussion of current algorithms and methods on redistricting, their strengths and weaknesses in term of performance, and factors considered especially on the geographical aspect. Then, the research expects a list of spatial redistricting requirements and specification according to the weaknesses in the existing methods.

After the research gains the requirement and specification, the research will contemplate a redistricting model by a working spatial redistricting prototype system based on the selected case study. The expected model will be integrated with advance knowledge and technology on concept of fuzzy logic, Multiple Criteria Decision Making technique and GIS environment. Subsequently, the prototype will consist of program modules, input and output. The program modules contain all the detailed processes of the proposed redistricting model with FMCDM. The program module helps to process the input and to generate optimal compact district for the district planner. Meanwhile, the inputs are from the user will provide the necessary data for redistricting and give the knowledge to the proposed prototype regarding to the compactness measurement.

Therefore, the research presents a metaphor, which provides a solution for automated and expert redistricting method. It is useful because it consolidate the most commonly occurring concepts that are the geographical concerns into a redistricting schema.

1.8 Significance of Expect Outcome

In essence, the model can contribute to different parties. The significance of this research study is toward the district planner, the governor and the public is elaborate as below:

1.8.1 Contribution for District Planner

The proposed model is obviously a great benefit for district planner. First, the process of redistricting becomes extremely fast and flexible because the data stored is in systematic way and redistricting process is done automatically. Furthermore, the proposed model helps redistrict by considering more than one criterion that it is more powerful compared to the conventional redistricting method, which commonly considers one criterion. Besides, the shape-based redistricting model redraws district boundaries with more compact district because the model concerns on spatial aspect, which often causes the less compact and discontinuous districts. This advantage not only saves the time of the district planners but also saves the work efficiency of the district planner. Moreover, human error will be minimized during redistricting process. In short, expected outcome will improve redistricting process in term of time efficiency, human intervention, and effectiveness.

1.8.2 Contribution for Governor

Redistricting with the proposed model will ease the governors who control, plan and use the land for many purposes in economy, politic and social. A more compact district plan will certainly help the governor to plan and manage the land resource more effectively. Less space will be wasted and all space will be utilized more precisely. In term of work force, governors need not to employ many workers or staff to conduct a redistricting process for election boundary redistricting. Meanwhile, transportation problem can be reduced for the school redistricting and governors will be able to minimize lost from transportation problem and overcrowding. Therefore, governors will certainly gain benefit from the proposed redistricting model.

1.8.3 Contribution for Public

The expected outcome from this research will contribute to the public. Public here refer to the different groups of people in various redistricting applications. For instance, voters are the public of election boundary redistricting and students and parents are the public of school redistricting. Conversely, the public will gain convenience compared to the governors. The voters will not suffer from transportation problem meanwhile students and parents will avoid overcrowding or transportation problem.

1.9 An Outline of the Dissertation

The outline of the dissertation is concurrent with methodology designed in section 1.7. For each different stage of the research, detail documentation is important to show the various results and issues of the research.

Chapter 2 describes the literature study of this research. Different types of redistricting applications show the usefulness and importance of redistricting process. Firstly, this research defines the various kinds of redistricting factors into two major categories: Application Independent factors and Application Independent factors. Then, this research presents and identifies previous algorithms and methods used in term of their backgrounds and capabilities. After that, this research compares different redistricting algorithms in order to show the strengths and weaknesses for each existing algorithm. On the other hand, this research concerns on methods to measuring compactness of the district plan. Existing compactness measurements are divided into two broad categories like Euclidean and non-Euclidean Measures. This research shows the increasingly important of compactness criteria in redistricting application. Lastly, this chapter mentions the specific problems, and the areas of the problem that will be solve.

Chapter 3 explains the development methodology. It contains an explanation on the solution and shows the general structure of the proposed shape-based redistricting model by using integration of Fuzzy Multiple Criteria Decision Making in the compactness measurement. Firstly, this research identifies the requirement specifications of the proposed model. Then, important tasks and domains are discussed before the explanation of the detail architectural design of the proposed model. Suitable compactness rules are selected to integrate appropriate compactness measurement methods with FMCDM to obtain an Integrated Compactness Indexing before it is used in the redistricting process. Finally, this research determines to use *location\allocation* and dynamic programming method in the redistricting process.

Chapter 4 has the detail discussion on the implementation of the prototype. This section discusses the development framework, Data Flow Diagram and the physical environment of the implementation process. Besides, there is explanation on the spatial data capture for the input data of the prototype. Then, different processes in the design stage in Chapter 3 will be implemented in the prototype by program coding. During the implementation process, this research conducted various testing such as module testing, integration testing, function testing and system testing to reduce the program error and to ensure the working of the prototype program.

Chapter 5 presents the result and analysis of the implementation. It explains detail evaluation of the developed shape-based redistricting algorithm using the prototype. Evaluation on the developed shape-based redistricting algorithm focuses on the applications and limitations of incorporating an integrated compactness indexing based on multiple compactness measurement by using Fuzzy Multiple Criteria Decision Making toward redistricting technique.

Chapter 6 discusses the research findings, accomplishments, limitations and future works on the developed shape-based redistricting algorithm based on the integration of multiple compactness measurements with Fuzzy Multiple Criteria Decision Making in a new compactness indexing after it had been conceptually designed, implemented, and evaluated. The main accomplishment of the research is the successful development of a compactness indexing and its use in redistricting algorithm to generate an optimal compact district plan. The achievement shows the contribution of the developed redistricting approach in improving redistricting technique.

CHAPTER 2: LITERATURE SURVEY

Ronald Reagan was not the only recent academic to state that computer could remove the controversy from redistricting and it can find the "optimal" redistricting plan by given any set of values that can be specified (Altman, 1998). Therefore, this chapter aims to present the literature survey on redistricting from different perspectives in aspects that appropriate to this thesis. First, the literature mentions the general categories of redistricting factors and their dependencies. Then, it justifies the status on the current two broad categories of redistricting methods. Among these methods, the review studies different kinds of algorithms and techniques in terms of their strengths, limitation, and the lacking in the redistricting process. Afterward, the review describes current measurements on the geographical criteria like compactness and continuity. More than thirty compactness measurement techniques or standards are listed and their current usability, consistency, and efficiency are discussed. This chapter also presents problem identification and analysis for providing an awareness of shape-based redistricting problem. These problems stimulate the proposed domain in the following chapter as a solution to the problem. In short, the outcomes of this chapter are to conduct reviews and analysis on all aspects of redistricting, and later to identify the specific problem of shape-based redistricting process for a better solution in an optimal compact district plan of existing applications.

2.1 Categories of Redistricting Factors

Factors considered in redistricting process are different among various redistricting applications like political boundary redistricting, school boundary redistricting and so on. However, this research divides redistricting criteria into two broad categories based on their dependency. The first category is application dependent criteria. These criteria may different between various kinds of redistricting applications and highly dependent on the purposes or goals of the application. The examples of these categories include population equality for election districts, workload balance for law enforcement district and potential client distribution for business district. The second category of the redistricting criteria is application independent criteria that are similar and unavoidable among different applications. These criteria refer to the geographical or spatial based redistricting that including the shape, size and boundary lines. The most common concerns for these criteria are continuity and compactness of the district. Continuity receives much attention in combination with compactness meanwhile compactness is attempts to capture the geographic regularity of districts. Both of the criteria are related to one another. For example, although the election boundary redistricting considers population equality for creating the election boundaries, a compact and continuous district is important too. Besides, there are techniques that weighting both of the criteria above together in the redistricting (Carmen, 2000). Consequently, these two criteria are highly related factors to create a fair and regular shape of district. Then, the higher concern for this research is on the spatial related criteria or the application independent factors without ignoring the application dependent criteria.

2.2 Advance Tools for Redistricting

With the increasingly importance of geographical criteria in redistricting, there is an important developments in the late 1980s contributed to a far greater use of computers for redistricting in the 1990s with the development of Geographic Information Systems (GIS) software. This particularly happened in United States in 1990 because almost every state in the United States used computers and GIS software as a tool for redistricting (Handley & Maley, 1997).

GIS is an useful computer-based system to collect, store, analyze, retrieve and display redistricting information because application independent criteria are important, related to and working together with application dependent criteria in the process of redistricting. The most distinguishing aspect of GIS compared to other Information System is its design to deal with objects and phenomena, where the geographic location in real world is critical to the analysis. Geographic data describes the real world in terms of their absolute location in real world based on coordinate system, their spatial interrelationships with each other, which describe how they are located to each other and their attributes, which are related to geographical, position (e.g. population distribution, demographic data, etc). Therefore, redistricting becomes much easier and manageable within the environment of GIS.

Minnesota was one of several states to use GIS for their election boundary redistricting on both congressional and state legislative districts around the year 1980. Other states using GIS for reapportionment following 1980 included Florida, Maryland, and New York (Hayes, 1997). Moreover, an automated redistricting system for Law Enforcement on an ARC/INFO based Police Beat and District realignment project is implemented in Montgomery County Police Department in Maryland (Fred & Monica, 1997). This system is more efficiently distribute its available law enforcement resources in response to the Community Policing Act and helps to create multiple redistricting recommendation. The system takes into account existing roads, natural barriers and various political features and based on the reported Police workload, develops redistricting recommendations for Police beats and Districts. Indeed, all these applications use GIS for the construction of redistricting models, the analysis of trends for redistricting data (e.g. demographic changes) and the visualization of the consequences of possible spatial decisions. Relationships among different types of data or coincidence of factors can be explored, such as finding areas with a certain combination of soil types and vegetation cover. In addition, there are even GIS software packages available that are specifically tailored for redistricting such as Maptitude® for Redistricting of Caliper Corporation's Maptitude GIS and Autobound.

GIS redistricting software provides the capabilities like to draw district plans interactively by assigning geography to districts piece by piece, produce maps, both on the computer screen and in printed form and generate statistical reports for the redistricting plan and each individual district in a plan. Moreover, some of the GIS software packages have the ability to customize and automate processing through their scripting languages. Each district is processed based on its unique shape, size, location, and whether it contained state boundaries.

In conclusion, redistricting nowadays is commonly using GIS advanced technology and it becomes an effective and efficient tool to simplify the redistricting process for all kinds of application. This technology helps to reduce the quantity of work force and minimize the cost for redistricting.

2.3 Redistricting Algorithm

2.3.1 General Redistricting Methods

There are many different applications for redistricting but some of the redistricting algorithms consider only the application dependent factors and therefore, they ignore the importance of geographical factors. On the other hand, most of the redistricting algorithms treat both these factors as a combination one. Thus, redistricting methods are used based on different type of redistricting factors and criteria. The criteria in redistricting process thereafter become the constraint of redistricting problems. Moreover, the complexity of the redistricting decisions causes the existence of many different methods in solving redistricting problems. Although there are many different methods, redistricting methods can be categorized into two broad categories: heuristic and exact methods. The section below explains these methods in detail especially on the status and usability.

2.3.1.1 Exact Methods

Exact method is a method used in redistricting and its descriptions are as below:

> **" *Exact* methods systematically examine all legal districts, either explicitly or implicitly. Explicit enumeration, or "brute force" search methods literally evaluate every district. More sophisticated methods include implicit-enumeration, branch-and-bound, or branch-and-cut techniques but exclude classes of solutions that can be inferred to be sub-optimal without an explicit examination. Finding the optimal districts in this case is then merely a matter of sorting the list of district scores." (Altman, 1998)**

From the definition, it is clear that exact method aims at finding optimal solutions and proving optimality. Thus, the advantage of this method is that it really provides an exact and optimal solution in a systematic ways. However, no exact method is developed for the time being that will solve redistricting problems for a reasonably sized plan and the mathematical structure of the redistricting problem makes it improbable to solve reasonably sized plans (Altman, 1997a).

2.3.1.2 Heuristic Search Methods

Heuristic procedures use a variety of methods to structure the search for high-valued redistricting plans. Most heuristics for locating optimal partitions are based on the principle of iterative improvement. In the field of computer science, heuristic algorithms are divided into two categories: *hill climbing* and *neighborhood search* techniques (Altman, 1997a). Firstly, the *hill climbing* methods work by making small improvements on a potential solution until a local optimum is reached. Hill climbing often starts with the current district plan or with a randomly generated plan and it makes improvements through repeatedly trading fully populated. The *neighborhood search* methods are all similar in that they seek to improve potential solutions by examining the value of "nearby" solutions; in this, they are similar to hill climbing. Unlike hill climbing algorithms, sophisticated techniques in this class use various techniques to attempt to avoid becoming stuck at local optima.

Most of the heuristic methods are used in political redistricting so that no one can expects the redistricting result and the political district plan will be fair with equal population distribution. Therefore, it appears to be popular method in this application but for other applications, heuristic methods do not promise to give an optimal solution. Besides, the modern meaning of heuristic methods is an algorithm that gets a "good" solution but not necessarily an optimal solution. Another commonly term used for this type of method is approximate method.

Heuristics are criteria, methods, or principles for deciding which among several alternative courses of action promises to be the most effective in order to achieve some goal (Sniedovich, 2000). This definition nicely captures the basic nature of many other definitions, such as, rule of thumb or a method not guaranteed to find the optimal solution but in practice often finds good quality solutions, where the term good is left suitably undefined. In other words, a heuristic is a search procedure that may give an optimal (best) solution to a problem but offers no guarantee of doing so.

Another way to look at heuristic here is a mechanism for guiding a search strategy by making intelligent short term guesses which are expected to result in the desired long term effect. Unfortunately there is no correct formal definition of what a heuristic is and heuristics encompass a wide range of methods from local search and genetic algorithms to Lagrangian relaxation (Sniedovich, 2000). The strength of heuristic or approximate method is that it can find solutions, in whatever problem size or characteristics. They provide often-good solutions but sometimes may loose time in trying to improve optimal solutions on small or medium size problems.

2.3.2 Comparative Discussion on Existing Algorithms

There are many different methods analyzed in automated redistricting for political purpose and therefore, the redistricting criteria include the population equality or/and geographical compactness. However, this research focuses the overall view in entire redistricting application, which is mentioned in Section 1.1. Redistricting also offers some interesting exercises in mathematics and computer science. In addition, algorithms for redistricting exploit techniques from computational geometry, graph theory, combinatorics and optimization methods (Hayes, 1997). This section discussed and analyzed different redistricting algorithms in respect of their the strengths and weaknesses.

2.3.2.1 Complex Decision making in Redistricting Process.

There are many decision making in redistricting process such as deciding the start point when redistrict, evaluating the optimal compact district and others. Therefore, this section describes some decision making problems faced and some necessity to improve redistricting process.

Redistricting process involved complex decision making. Firstly, redistricting problem may involve more than one criterion for both application dependent and application independent criteria. For examples, redistricting without considering the geographical compactness will certainly create the worst district plan in term of district shape (Carmen, 2000). Furthermore, the redistricting decision is a difficult one because the multiple criteria may depend on one another to influence the drawing of the district boundary. Moreover, there are certainly some Differences on the importance of different criteria in affecting the result of a redistricting application. For instances, for redrawing school district, the district planner may need to consider student distribution or population (most important), the nearby housing estates (less important) and the capability of the school itself (important). In short, redistricting involves many decisions making in order to achieve an optimal result and involve of many criteria (Altman, 1998). Besides, redistricting decision will cause many effects to different application such as the school redistricting will affect on house values and redistricting in business may cause the success or the failure of a particular business.

Most of the redistricting uses relative method to compare the district criteria (Carmen, 2000; Altman, 1997; Clark et.al., 1999). The relative value here may means the equal relative population or population equality in election boundary redistricting for good results. However, the decision making and the comparison between relative population and even shape compactness are in vague and imprecise condition. For examples, in State of Texas, each district in a plan that is drawn by a redistricting process must comply with criteria that have a population that is 'close to' the population of all other districts in the plan (Clark et.al., 1999). The meaning of 'close to' here shows a vague value to indicate an ideal district.

Another impediment to the use of many optimization methods is the degree to which the results hinge on start point (Carmen, 2000). The decision on the selecting the start point is an important stage, which may influence the result of the redistricting plans. For instance, Carmen (2000) had found the major drawback to use of optimization methods like *simulated annealing* and *location\allocation* that they look at situations on the efficient new field of study whereas there is no consideration of the interior point (Carmen, 2000). Carmen (2000) attempted to use the same approach for a election boundary redistricting for congressional redistricting in Louisiana, but terminated the program after 5 days without optimum solution because the method was able to generate 20,000 plans within 1 day.

Therefore, this research will aim to improve the decision making process in redistricting. This research will treat redistricting as a multiple criteria problem. Besides, the research will use the theory of fuzzy logic in decision making to create a suitable context for district planners to evaluate the district shape compactness and to draw district boundary. On the other hand, fuzzy logic allows decision making with estimated value under incomplete and uncertain information (Kashani, 1996). Besides, it breaks the redistricting decision scenario into small parts that one can focus and then combines the scenario pieces into a global interrelated whole so that one can indicates which alternative is the best (produce the decision (choose the alternative) that best meets the criteria).

2.3.2.2 Implementation Result for Existing Redistricting Algorithm

Redistricting was implicitly characterized as a combinatorial optimization problem by Vickrey (1961) onward. Blevins (1999) suggested that political redistricting is an optimization problem and therefore optimization algorithms like integer optimization may solve it. These optimization algorithms are used for school redistricting.

Optimization provides a means of determining the best answer to a given set of circumstances. For instances, in election boundary redistricting application, optimization here usually means to determine the best answer for the population equality. One of the optimization solutions used in redistricting application is *cake-cutting* method that based on the theory of fair division. Fair division here is achieved by finding ways to obtain envy-free divisions. Similarly, it is always possible to fairly divide a cake among *n* people using only vertical cuts. Furthermore, it is possible to cut and divide a cake such that each person believes that everyone receives $1/n$ of the cake according to his own measure (Weisstein & Stephen, 1999). However, the weakness for this method is that the details in this method are crucial although decisions of this method may be changeable and have to be stated explicitly. Furthermore, the trickiest part for this method is for maintaining the continuity. This procedure can run into blind path, leaving a district stuck to others which are already allocated and cannot create contiguous districts (Hayes, 1997).

Besides *cake-cutting* algorithm, *greedy* algorithm is another possible optimization solution for redistricting. The idea behind a *greedy* algorithm is that it produces a good solution quickly by selecting a shortsighted best addition until a set is constructed that is feasible (Greenberg, 1998). It is an algorithm used to recursively constructs a set of objects from the smallest possible elemental parts (Weisstein & Stephen, 1999). The algorithm is greedy because it always examines the "next best" with no backtracking. It is a maximal solution in the sense that no solution that includes the selected items has a better return. This algorithm always picks the highest portion of area that contiguous to its own region. Therefore, the potential trouble spot in *greedy* algorithm is to get the procedure start. Then, it is easy to reflect pathological cases where all the largest districts are in tight cluster, producing seriously ill made district (Hayes, 1997).

Then, *steepest ascent* algorithm is to improve the *greedy algorithm*. *Steepest ascent* means to maximize while *steepest descent* is for minimizing (Greenberg, 1999). This is a class of algorithms, for finding the nearest local minimum of a function (the smallest value of a set or function), which assume that the gradient of the function can be computed. The steepest descent method starts at a point P_0 and, as many times as needed, moves from P_i to P_{i+1} by minimizing along the line extending from P_i in the direction of the local downhill Gradient. (Weisstein & Stephen, 1999). It has post-processing stage to optimize the alignments of *greedy* method. However, it is easily being trapped in a local optimum. Besides, the weakness of such schemes is that the marble can stuck in a local trough and never find the optimum lowest point (Hayes, 1997). Therefore, this technique needs mechanism for escaping from local minimal.

Simulated Annealing algorithm is an improvement after *steepest ascent* algorithm. It is a common computer science approach that uses some randomness to "shake things up" and then lets them "settle down" into a better redistricting. *Simulated annealing* is one of the most successful of combinatorial optimization methods (Altman, 1998). It is based on a mathematical analogy to the slow cooling of metal. If the value function being optimized is sufficiently well behaved, *simulated annealing* asymptotically converges to the optimum value. It was previously recommended for use in redistricting (Browdy, 1990). This algorithm then is to overcome problem of the *steepest-descent* algorithm. It allows occasional detrimental (disturbing) moves. The probability of accepting an unfavorable move is determined by a parameter analogous to a temperature, and the system is "annealed"(strengthened) by steadily reducing the temperature toward zero where only favorable move are possible. However, this algorithm is not deterministic (rational) because the randomness will causes repeated runs on the same input data yield different output (Hayes, 1997). For example, Carmen (2000) had recently applied this approach and was able to generate 20,000 plans within one day. However, this technique has generally employed large aggregate units such as counties in election boundary redistricting and has often combined countries to reduce the complexity of the problem (Mehrotra et. al., 1998; Carmen, 2000).

An implicit enumeration technique is one of the redistricting methods mentioned by Liittschwager and colleagues (1973) for Iowa Senate Districts. Indeed, Balas (1965) gives a brief description of these methods:

> **"Enumerative (*branch and bound, implicit enumeration*) methods solve a discrete optimization problem by breaking up its feasible set into successively smaller subsets, calculating bounds on the objective function value over each subset, and using them to discard certain subsets from further consideration. The bounds are obtained by replacing the problem over a given subset, with an easier (relaxed) problem, such that the solution value of the latter bounds that of the former. The procedure ends when each subset has either been reduced to a feasible solution, or has been shown to contain no better solution than the one already in hand. The best solution found during he procedure is a global optimum."**

However, these methods are usually very sensitive to the goal functions. It is important to know quite a lot about the goal function to derive the "relaxed" problem in order to use to set bounds. The restrict input set to guarantee that the branch-and-bound procedure does not take exponential time (Altman, 1997b).

Lastly, *Genetic algorithms* are search algorithms based on an analogy to natural selection and genetic combination. Potential solutions to the optimization problem are defined as genetic strings, which can be mutated or "crossed" with other strings (Altman, 1998). A group of potential solutions then competes to survive and reproduce in the next generation. It requires the specification of three operations (each is typically probabilistic) on objects, called "strings" (these could be real-valued vectors) (Greenberg, 1999):

Reproduction - combining strings in the population to create a new string;
Mutation - spontaneous alteration of characters in a string;
Crossover - combining strings to exchange values, creating new strings in their place.

Unlike other subsymbolic search algorithms, like *simulated annealing*, the probabilistic primitives that *genetic algorithms* use to manipulate their population and their lack of an explicit memory make them very fast on contemporary hardware (Rawlin, 1998). Various theoretical results on *genetic algorithms* have been proved that are of fundamental significance to their operation and much is known about their basic behavior. However, there are few formal shows the results about the their tradeoffs.

In brief, redistricting algorithm by using optimization solution is well established because there are many algorithms used to solve redistricting problem. From the analysis above, it is clear that all kinds of optimization solution have their strengths and weaknesses. Therefore, this research aims to apply optimization technique to solve redistricting problem in order to improve the existing redistricting solution. The output from this research will be the optimal compact district plan.

2.3.2.3 Redistricting is NP-complete

Later, there are many researches in the area of computer science and operational research have examined redistricting problem and reached some conclusions about their computational complexity. Moreover, Altman (1998) showed that the redistricting problems are likely to be intractable. Therefore, redistricting is categorized as a NP-complete problem, which is non-deterministic polynomial (Weisstein & Stephen, 1999). Non-deterministic here means that no particular rules is followed to make the guess. If a problem is NP and all other NP problems are polynomial-time reducible to it, the problem is NP-complete. Thus, finding an efficient algorithm for any NP-complete problem implies that an efficient algorithm can be found for all such problems, since any problem belongs to this class can be recast into any other member of the class. It is not known whether any polynomial-time algorithms will be found for NP-complete problems, and determining whether these problems are tractable or intractable remains one of the most important questions in theoretical computer science. One approach when a NP-complete problem must be solved is to use a polynomial algorithm to approximate the solution. However, the solution thus obtained will not necessarily be optimal but will be reasonably close.

Nevertheless, two algorithms that can be used to solve redistricting problems as NP-complete problems are fully polynomial algorithm and approximation algorithms. First and the former, computational complexity is a measure of the difficulty of obtaining optimal solutions, but it says little about the difficulty of approximation. Some problems are much easier to approximate than others are. For some NP-complete optimization problems, methods in polynomial time generate a solution that is guaranteed to be within a particular percentage of the optimal value.

Methods that obtain arbitrarily close "solutions" in polynomial time are known as fully polynomial approximations. Unfortunately, it can be proved that no fully polynomial approximations exist for many of the redistricting sub-problems discussed above. Although the redistricting problem does not allow arbitrarily close approximations, the research has no excluded the possibility of approximating a solution to within some fixed percentage the optimum. No such guaranteed approximation procedure for a redistricting sub-problem is yet been demonstrated, but the question remains open.

Then, approximation algorithm is algorithm guaranteed to find close to optimal solutions. For example, the travelling salesperson problem in the plane is NP-complete, but in polynomial time it can be solved approximately within every constant. Approximation algorithms have developed in response to the impossibility of solving a good many problems exactly (Cowen, 1998). Trading-off optimality in favor of tractability is the paradigm of approximation algorithms.

Although redistricting is proved to be NP-complete problem but the suggested algorithms above still do not have clear and systematic examination on the effectiveness for redistricting problem. Therefore, this research will not use the suggested solution for redistricting problem. In other words, this research will treat redistricting as multiple criteria decision problem and optimization problem. This has lead to the solution by using the method from these two domains.

2.4 Geographical Concern for Redistricting: Compactness and Continuity

Compactness of a district is determined by considering its appearance and the area of dispersal of the district (Altman, 1998). This should include a mathematical analysis of the shape and regularity of the spatial shape. Therefore, irregular geographical boundaries may cause oddness in district shape. In addition, Laurini & Thompson, (1992) mentioned that shape refers to the structure of the mode of arrangement related to function or the appropriateness and effectiveness of purpose. Subsequently, compactness is important in control of the shape of districts and district boundary complexity or compactness interdependent to district shape.

Generally, compactness measurement techniques can be summarized as in Figure 2.1. They are categorized as Euclidean and non-Euclidean measures. Euclidean measure uses Euclidean parameters like area, perimeter, width, and length. On the other hand, non-Euclidean measure does not use such parameters, like fractal analysis that uses the fractal dimension to measure the district line complexity and other application dependent criteria. Although there are different measurements, each measurement aims to assess district compactness and continuity with certain useful parameter. The example of the fractal dimension is useful to assess the district compactness and continuity through the fractal analysis on district boundary complexity. There are many different methods like structure walk, box-counting, dilation and Euclidean Distance Mapping (EDM) that can be used to get fractal dimension (Dominique & Michel, 1999). However, these methods react or perform differently with different considerations such as district size, dimension, pixel and techniques of implementation.

Various types of compactness measurement for evaluating the result of the district plan are from tremendous endeavor of compactness measurements in redistricting. It really shows the importance of geographical aspect in redistricting and necessity for different techniques to be used to measure the compactness and continuity of the district plans. However, these compactness measurements commonly used after the redistricting process and usually treated as separately to redistricting effort. The reason for assessing compactness and continuity of the redistricting plan is to ensure balance and fair result (Carmen, 2000).

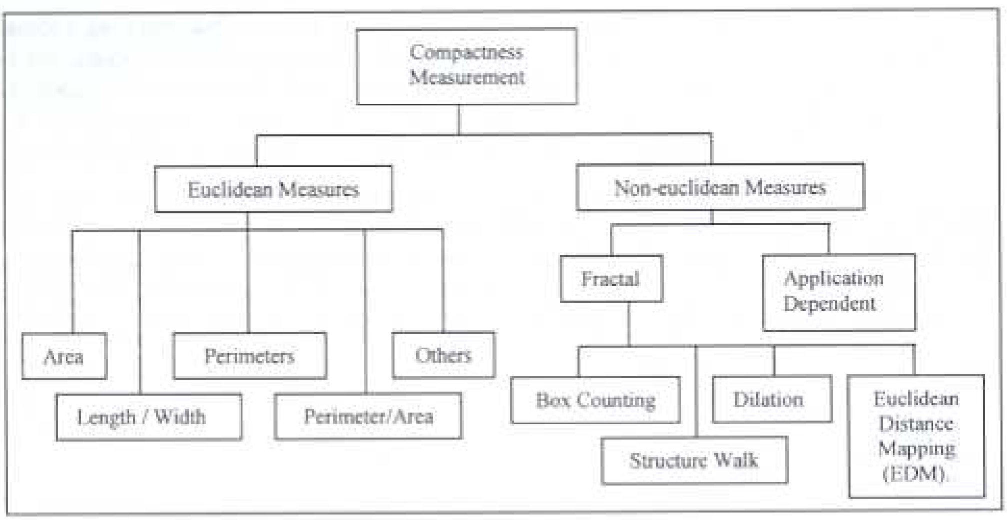


Figure 2.1: Summary for type of Compactness Measurement

2.4.1 Euclidean Compactness Measures

Recent surveys of compactness criteria list more than thirty different measurement formulas. Altman, (1998) shows the different type of compactness measures and most of the compactness measures are characterized by Euclidean Measures. Some of them are parametric representation. On the other hand, according to Laurini & Thompson (1992), shape and compactness measures may utilize sets of parameters defining particular properties of areal objects or may use data for sets of line segments fitted to the boundaries. Therefore, this research summarizes these Euclidean compactness measurements as the four tables below based on the parameter used. The four groups of parameters used including the length and width, area, ratio of perimeter with area and others (Altman, 1998).

Table 2.1: Length v. Width Earliest Use

LW 1: W/L: where L is longest diameter and W is the maximum diameter perpendicular to L
LW 2: W/L: from circumscribing rectangle with minimum perimeter
LW 3: 1/(W/L): rectangle enclosing district and touching it on all four sides for which ratio of length to width
LW 4:W/L, where L is longest axis and W and L are that of a rectangle enclosing district and touching it on all four sides
LW 5: L–W where L and W are measured on north south and east-west axes, respectively
LW 6: Diameter of inscribed circle/diameter of circumscribed circle
LW 7: Minimum shape diameter/maximum shape diameter

Table 2.2 Measurements Based on Area

A 1: The ratio of the district area to area of minimum circumscribing circle
A 2: The ratio of district area to the area of the minimum circumscribing hexagon 32
A 3: The ratio of district area to the area of the minimum convex shape that completely contains the district
A 4: The ratio of district area to area of the circle with diameter equal to the districts' longest axis
A 5: The area of the inscribed circle/area of circumscribed circle
A 6: The area of the inscribed circle/area of shape
A 7: (area of intersection of the shape and circle of equal area)/(area of the union of the shape and the circle of equal area)

Table 2.3: Measurements Based on Perimeter/Area Ratios

PA 1: The ratio of district area to the area of circle with same perimeter
PA 2: $1\text{-}PA_1^{1/2}$
PA 3: The ratio of perimeter of the district to the perimeter of a circle with equal area
PA 4: The perimeter of a district as a percentage of the minimum perimeter enclosing that area (=100(PA_3))
PA 5: A/0.282P
PA 6: $A/(0.282P)^2$

Table 2.4: Other Shape Measures

OS 1 :	The moment of inertia — the variance of distances from all points in the district to the district's areal center of gravity, normalized. Where A is the area of the shape, r is the distance from the center and D is the set of points in the shape this is $\frac{A}{\sqrt{2\int D \int r^2 dD}}$
OS 2 :	The average distance from the district's areal center to the point on district perimeter reached by a set of equally spaced lines
OS 3 :	(radius of circle having same area as shape)/(radius of circumscribing circle)
OS 4 :	(N-R)/(N+R) where N,R is # of (non)reflexive interior angles (respectively)

2.4.2 Non-Euclidean Compactness Measurement

Non-Euclidean compactness measurement refers to the measurement that does not use Euclidean features in the measures. They include the application-dependent compactness measurement and fractal analysis as discussed as following.

2.4.2.1 Application-dependent Compactness Measurement

Among various kind of redistricting applications, political redistricting creates high popularity in most of the research studies (Carmen, 2000; Altman, 1997b). Therefore, among various kinds of application dependent criteria, population distribution becomes the most commonly uses for compactness measurement. The tables below show some of the population-based compactness measurement and plan compactness measures. If the population density is uniform in political redistricting application, the population-based measurements are equivalent to the geographical measures A3, A1 and OS1 respectively in the previous section (Altman, 1998).

Table 2.5: Population-based compactness measures

POP1:	Ratio of district population to the population of the minimum convex shape that completely contain the district.
POP2:	Ratio of district population to the population in the minimum circumscribing circle
POP3:	Population moment of inertia, normalized
POP4:	Sum of all pair-wise distances between centers of subunits of legislative population, weighted by subunits of population

Table 2.6: Plan compactness measures

PL1:	The sum of district perimeters
PL2:	The maximum absolute deviation from the average district

2.4.2.2 Compactness Measures Using Fractal Analysis

The application of fractal analysis especially in GIS environment provides additional tools to gain insight into the spatial nature of compactness (Knight, 1997). Fractal concepts are one of the techniques that can be applied in geographic study to evaluate spatial characteristics especially in redistricting. Geographic shapes, such as voting districts, are often complex and have been described by the media, public and the courts as "serpentine", "growing bacteria", "spiders", "shape with tortuous nooks and crannies", and "DNA fragments" (Shiode, 1998; Knight, 1997). All these descriptors are used to relate shape characteristics. The expressive language characterizes geographic shapes as complex in nature. The Random House Dictionary of the English Language (1987) defines a fractal as:

> **"A geometrical or physical structure having an irregular or fragmented shape at all scales of measurement... such that certain mathematical or physical properties of the structure... behave as if the dimensions of the structure are greater than the spatial dimension."**

Fractal geometry is a tool to describe erratic, complex forms of nature which are neither points, lines, or areas but instead fall between the traditional categories of topological dimensions (Muller, 1986). A fractal dimension D, which will fall between one and two (assuming that no geographic polygon is so small that it becomes a point), can describe the complex nature of geographic shape. A fractal dimension allows discussion of values between one, two orthree dimensions as summarized in following table.

Table 2.7: Geographical features with different topological dimensions

Geographical features	Topological Dimension
Point	0
Boundary (Curve)	1
Surface	2
Solid	3

Although different geographical features have different topological dimensions, a complex curve may wander on a surface. In the extreme case, the curve may be so complex that it effectively fills the surface on which it lies. A curve will have a fractal dimension of a real number between one and two. A complex curve that approaches surface filling will have a fractal dimension approaching two. Therefore, the more complex the geographic boundary (curve), the higher the fractal dimension. Boundaries will have a fractal dimension somewhere between a straight line, with a fractal dimension of one, and a boundary that is so complex it is space filling with a fractal dimension of two. Fractal dimension can be used as a model for measuring the complexity of the district boundary (curve) or district's area (polygon) (DeCola, 1993). The complexity of geographic areas can be described by estimating the fractal dimension (D).

Fractal analysis (fractal dimension) may provide a clear approach to analyze the spatial structure and form than current compactness measures. It appears feasible and reasonable to apply fractal analysis to define district complexity or compactness especially with fractal figures which share the three types of features (Shiode, 1998). Firstly, fractal figures are non-differentiable. They are different than geometric shapes, which usually have definite scales and lengths such as radius or circumference of a circle and the edge or diagonal of a square to characterize themselves. Secondly, fractal figures hold self-similarity, their shapes do not change even when observed under different scale. This nature is called scale-invariance. Even the geo-data used in GIS, such as coastline shape and urban space pattern, hold strong statistical self-similarity between different scales. Therefore, when zoom in closer to a landscape image and as the level of detail increases, a pattern similarly will be found at a larger scale level. Finally, the actual value of these fractal figures differs slightly, depending on the method of defining it. Some of the known techniques used for calculating the fractal dimension are structure walk, box-counting, dilation and Euclidean Distance Mapping (EDM) are discussed in Appendix A.4.

The use of fractal dimension to evaluate the complexity of shape boundaries has the potential of providing the advantages, which are not found in conventional compactness measures. Indeed, fractal dimension appears to be superior in evaluating the full range of boundary complexity with the ability to summarize the district complexity (Shiode, 1998; Knight, 1997). Conventional compactness measures use very specific and limited geometric parameters to represent the whole district (Shiode, 1998; Knight, 1997). This allows specific shape characteristics to unduly influence the resulting measure. Besides, all conventional measures must consider the geographic shape as a whole. If a geographic area contains a meandering river or oddly shaped state boundaries, it still must be included in conventional measure calculation. Consequently, the measures may be highly influenced by factors that are not included as part of the intended evaluation. Not all fractal measures require the boundary be a closed polygon. This allows the focus of the investigation to be modified. Geographic boundaries that are not subject to study, such as the state boundary or shorelines, may be removed.

2.4.3 Current Status of Shape-based Redistricting

Compactness measurement for district compactness and continuity assessment gives an indexing for assessment of compact district. Frequently, decision making in this district assessment involves uncertainty of the data acquired and the variety of evaluation tools. In addition, numerical measurements in compactness measurements give uncertain and vague value for accurate assessment of district compactness and continuity. Altman's axiomatic assessment in Appendix A.1 on the compactness measurements shows that existing compactness measurements do not accurately assess compact district shape. Furthermore, Dominique & Michel (1999) showed different methods for getting fractal dimension may change due to different condition. Thus, existing compactness measurement methods performs differently to different circumstances and single methods alone is not effective enough because each of the method may suffer from its limitation.

There is a lot of redistricting techniques but still does not give an ideal district plan in term of geographical aspect. Redistricting cannot run away from compactness and continuity because many redistricting problems as discussed in Section 2.3 are suffering for the spatial context problem. However, there is no incorporation of compactness measurement to redistricting techniques to get optimal compact district. Shape-based redistricting in this research will ensure the compactness and continuity in the redistricting process in order to obtain an optimal compact district plan in term of the geographical aspect. This will surely contribute to the governor, district planner and public to avoid the development of a not practical district plan which caused by the geographical factors.

Next, shape-based redistricting is a multiple criteria problem because redistricting must be done with at least an application dependent criterion and together with application independent criteria such as continuity and compactness. Thus, multiple criteria will be considered to draw a district plan in this research. General redistricting, which aims to maintain optimal compactness, continuity and other application dependent criteria is a complex multiple criteria problem because minor changes in any of the criteria may affect the other criteria.

In study of Euclidean compactness measurement, Altman (1998) uses an axiomatic analysis to join formal measures of compactness to analyze the consistency of these compactness criteria. These definitions can help formally define what it means a compact measure to capture shape in six axioms. The detail of six Altman's axioms is discussed in Appendix A. The evaluation from the experimental results of Altman's axioms shows that most of the Euclidean measures suffered from inconsistencies. Most compactness indices reflect at least one principle of shape manipulation. However, Altman (1998) shows that these measures satisfy one shape axiom, but violate others. The shapes are not generally rectangular, straight or circular and cannot be represented by a simple analytical function. They have a more complex nature that is not easily defined. Subsequently, the seemingly simple task of describing a complex shape is often an intimidating task.

2.5 Conclusion

The redistricting methods and algorithms are extremely important in order to draw a fair and unbiased district. District plan is defined as the output of the redistricting process and the standards to evaluate the district plan refer to the compactness measurement techniques. Therefore, this chapter covers a discussion on the redistricting backgrounds, complexity, and problem of redistricting algorithms and compactness measurement. The existing redistricting algorithm shows their strengths and abilities but they also bring weaknesses in creating an ideal district plan and even face difficulty in implementation process. It is clear that redistricting is a multiple criteria problem and there is a need to improve existing algorithm especially in term of geographical aspects. The compactness measurements encompass continuity and therefore it means that the maximally compact plan will contiguous (Altman, 1998). Although the assessment tools are called compactness measurements, they assess not only on the compactness but also on the continuity of the districts in the plan. Therefore, the proposed redistricting model will incorporate compactness measurement to determine the shape compactness rather than to use the measurement method only after the redistricting process. Consequently, the proposed model will be a shape-based redistricting because it considers redistricting in spatial perspectives. Therefore, the outcome of this chapter enables the research to identify general requirement of the proposed model. The proposed model shall be able to handle decision making for an improved compactness measurement indexing. The decision making method will be more descriptive and able to incorporate with natural feelings like their confidence and their attitude to risk of district planners or decision makers. During decision making, it should be cope with fuzziness and multiple criteria. Therefore, the model will integrate the strengths of particular method and at the same time reduce or minimize its weaknesses or lacks. Lastly, the proposed model shall work and perform in an environment (software or hardware interface) that is able to manage the spatial and non-spatial data, their relationship and dependency.

CHAPTER 3: CONCEPTUAL DESIGN OF INTEGRATED COMPACTNESS INDEXING FOR SHAPE-BASED REDISTRICTING

After the literature and analysis from the previous section, a model is proposed specially to solve redistricting problem according to the optimization and multiple criteria decision making method. The proposed model is by using Fuzzy Multiple Criteria Decision Making (FMCDM) through Dynamic Programming (DP) approach for enhancing the shape-based redistricting in term of the decision making and optimization of districts shape compactness. Therefore, compactness measurements are used as the main criteria for the shape optimization in the proposed model but the description for the shape of district is usually linguistics and vague. Moreover, there are multiple standard and methods in the shape evaluation. Thus, Multiple Criteria Decision Making (MCDM) is a suitable solution because it enables the consideration of the multiple aspects of compactness methods with relative weights for their importance in the shape-based redistricting. Furthermore, Fuzzy Set Theory (FST) allows the consistent and systematic decision making in vagueness and subjectiveness of redistricting environment. Therefore, FST and MCDM in FMCDM will be used to generate a new and integrated compactness measurement index that considers multiple spatial and non-spatial factors or criteria. Then, this research uses the integrated index together with the knowledge of DP and location/allocation to conduct the shape-based redistricting. Consequently, shape-based redistricting optimization and multiple criteria problem will be solved directly with the improvement in the new index and its application in shape-based redistricting.

This chapter describes the design of the proposed model for shape-based redistricting by an improved compactness indexing. Requirement specifications are identified before the discussion on the task and domains of the architecture design for the proposed model. Therefore, the architectural design of the model will be described followed by the detail discussion on the proposed model. The detail discussion covers the explanation on the data preparation, knowledge acquisition, techniques used for the FMCDM and DP to provide more understanding on the methods selected. This study discusses the illustration of the general schema of the FMCDM to a case study in order to visualize the model functionality and capability for the techniques.

3.1 Requirement Specification

The proposed model aims to get optimum good or compact district, which defined as in this research is as regular and contiguous as possible. The main significance is to ease the control of a space. For example, with regular and continuous districts, governors face less difficulty in ruling a district or even country in political redistricting. On one hand, it helps to pay concern on the minority group for any available facilities or services in a district. Besides, in almost every other single applications mentioned in Section 1.2, compact districts enable accessibility of all services and facilities, avoid overcrowding cases, help to minimize transportation costs and addressing impact on high growth areas. Specific requirements for the proposed solution is as the following:

(1) To integrate multiple compactness measurements to gather the strengths of particular method and at the same time to reduce or minimize its weaknesses or lacking.
(2) To produce better or integrated shape assessment indexing that is more descriptive and able to incorporate with natural feelings of district planners. Natural feelings here include their confidence and their attitude to risk.
(3) To cope with fuzziness in the shape assessment indexing.
(4) To incorporate the new indexing into redistricting algorithm to generate an optimal compact district.
(5) To consider the restricted boundary like river, political boundary and any others during the redistricting process.
(6) To work and perform in an environment (software or hardware interface) that is able to manage the spatial and non-spatial data, their relationship and dependency.

In order to get optimum shape in redistricting, compactness measurement will be necessary to measure the regularity of the district shape. Therefore, the research proposes to use these compactness measurements as goals to draw district plan instead of using them to assess the compactness after redistricting process. Multiple compactness measurements will be chosen to help to assess the district compactness and continuity because one compactness measurement alone may not be a helpful and promising assessment tool. Selecting multiple compactness measurement enables the evaluation of the district boundary complexity and the analysis of relative district compactness in a district plan. At the same time, compactness measurement also tends to measure the continuity of the district and non-contiguous district will be classified as not compact district in this research.

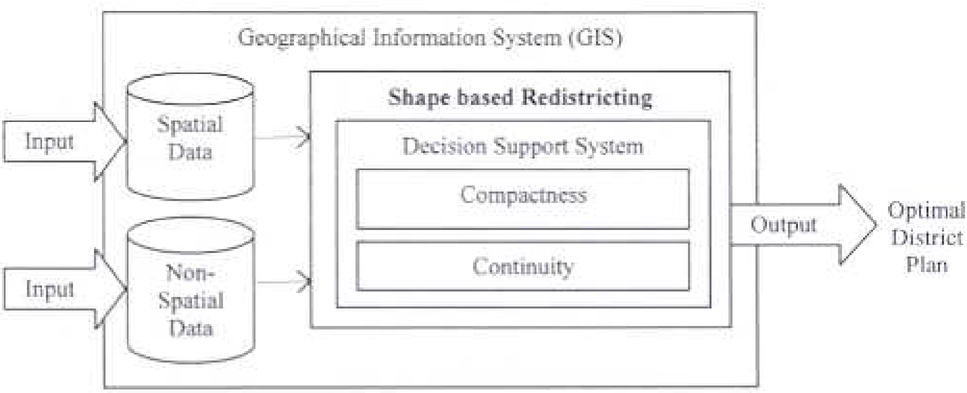


Figure 3.1: Revised redistricting model for shape-based redistricting

3.2 Task and Domain

The specific requirements in previous section stimulate the proposed task and domain in this section. The domain for this research is for shape-based redistricting. Figure 3.1 briefly shows the revised shape-based redistricting model in terms of the input, process, and output. Thus, the spatial data is the most important input data and other data components are the related non-spatial data. The task of the proposed model is an enhancement of the decision support system in maintaining the compactness and continuity of the redistricting to produce the optimal compact district plan (the output).

This proposed model considers multiple compactness measurements method to ensure optimality of compact districts. However, the compactness measurement provides numerical indexing, which is vague and may be incomplete. Therefore, this study will use Fuzzy Multiple Criteria Decision Making or namely FMCDM method for the solution. Multiple Criteria Decision Making (MCDM) deals with problem of helping the decision maker to choose the best alternatives based on several criteria (Fuller & Carlsson, 1996). This approach helps to solve complex shape-based redistricting problems in a systematic, consistent and way that is more productive because it enhances the degree of conformity and coherence in the decision process by giving a set of optimality rules. By then, the model able to gain a compactness assessment indexing, that is the decision criterion and the statement concerning the objectives of the problem. On the other hand, the fuzziness concept may helps to solve the subjectivity and vagueness for deciding the compactness assessment indexing. Therefore, this research with fuzzy set theory in MCDM model provides an effective way for dealing with the subjectivity and vagueness of decision making process for multiple criteria shape-based redistricting problems. Subsequently, FMCDM method supports a systematic decision making for several reasons (Zimmermann, 1995). First, the information and knowledge for the redistricting decisions is incomplete, uncertain, imprecise, or even inconsistent. Second, there are multiple conflicting goals and multiple different type of constraints in shape-based redistricting. FMCDM which is an Operation Research (OR) technology (Herrera & Verdegay, 1995) able to face the complexity of redistricting problems and to provide needed strategic decisions.

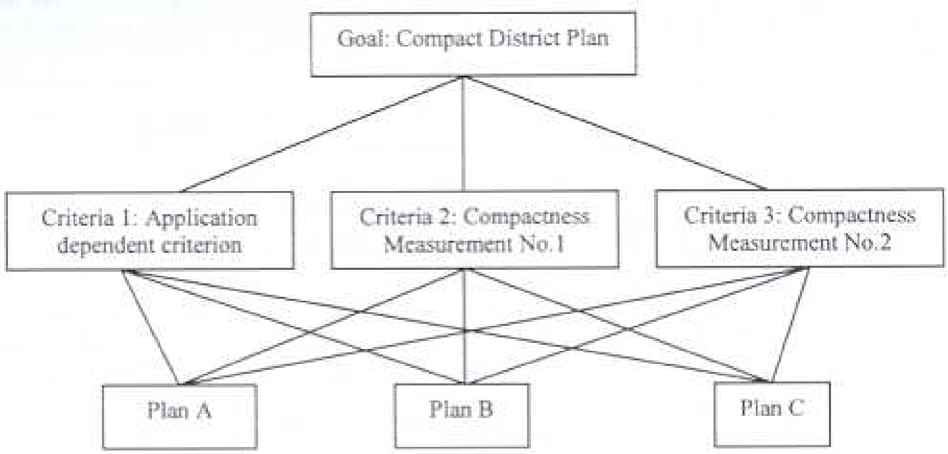


Figure 3.2: Hierachy representation of the problem

Furthermore, this research able to solve the redistricting problem specifically concerns on the geographical aspect with the help of FMCDM. Hierarchical representation in Figure 3.2 shows the goals, multiple criteria, and alternatives according to the FMCDM methods. The goal for the shape-based redistricting is for optimal compact district plan. Besides, this research will use two criteria on compactness measurement so that the geographical factor could helps to determine the optimality of the shape compactness. An another application dependent criterion is used to help the working of redistricting process for any specific application. For instance, this criterion is on population for election boundary redistricting and it is on student distribution for school redistricting. This application dependent criterion ensures the proposed model to be a general model that can be used in any application. Therefore, the output of the district plan not only concerns on the application factor but also concerns on the geographical aspect to prevent any not compact or not practical district shape. In other words, the study will accomplish the specification (1) to (3) in previous section.

In the multiple criteria program, redistricting application functions are established by measuring the degree of fulfillment of the decision maker's requirements about the goal function and are extensively used in the process of finding "good compromise" solution (Fuller & Carlsson, 1994). The requirements in redistricting application include the achievement of goals on compactness, nearness to an ideal point on the application dependent factor, and other satisfaction. Fuller and Carlsson, (1994) mentioned that one of the earliest practical application of FMCDM is a commercial application for evaluation of the credit-worthiness of credit card applicants. Besides, the recent applications including the evaluation of weapons systems, a project maturity evaluation system at Mercedes-Benz in Germany, technology transfer strategy selection in biotechnology and aggregation of market research data.

Then, the adopted method in this research is Dynamic Programming(DP) method which is commonly used to solve the knapsack problem. In other word, this study treats redistricting problems as similar to a knapsack optimization problem, which is maximization or minimization of a value. For example, a thief robbing a safe finds it filled with N items of various size and value but bag has only limited capacity (M). In contrast, getting compact district plan means to calculate the best combination of individual district for all district size up to total district plan size. The study applies DP method based on two main reasons for the specification (4) in the previous section. First, DP can build a good model in a bottom-up technique to solve redistricting problems. It allows the breaking up of all problems into a sequence of easier subproblems which are then evaluated by stages and has the power to determine the optimal solution by solving each of these stages optimality (Render, 1997). Second, redistricting decisions is usually accompanied with many complicated considerations, so these might be numerical constraints or some constraints which were the experience of the experts. These constraints are difficult to solve by nonlinear programming or other methods but can be easily incorporated and solved by the DP method. In addition, when integrates with FMCDM, these constraints can even be systematically analyzed and solved.

This research exploits the uses of GIS software to smooth the processing of the raw data to the proposed model. The use of GIS is according to the specification (5) because GIS enables easier way to maintain the redistricting data and helps to maintain the relationships among the data. In addition, this technology provides ready-to-use geometry data such as area or perimeter for each of the district. According to the complex functions in shape-based redistricting, GIS provides easy-to-use and extended functions like grid analysis, network analysis and overlay analysis which are extremely helpful for each of the redistricting process in this proposed model in two ways. Firstly, GIS enable modeling of real-world features with points, lines, and areas with symbols and labels (text) that describe objects. For instances, points define discrete locations of geographic features and represent locations that have no area, such as mountain peaks. Lines represent the shapes of geographic objects, which are too narrow to depict as areas, such as streets and streams, or linear features that have length but no area, such as elevation contours. Then, areas are closed figures that represent the shape and location of homogeneous features such as states, counties, and parcels. Secondly, topology in GIS explicitly defines spatial relationships such as a polygon is defined by the list of arcs comprising its border that like a district with its boundary. Topological relationships enable data to be stored efficiently, so large data sets can be processed quickly and it facilitates analytical functions such as identifying adjacent district features and so on.

In short, this research proposes a method for shape-based redistricting by the FMCDM methods in compactness measurement to acquire optimal compact district plan. The compactness measurement technique is used to measure the shape compactness or regularity and continuity during redistricting process so that the violation of each of the compactness measurement method can be reduced (refer Section 3.2.4). Besides, the advanced OR technique with the fuzzy set theory will be helpful in vague compactness measurement indexing. Then, Dynamic Programming will be promising for the optimality in the proposed model because this method simplifies the problem by dividing a complex redistricting problem into sub-problems, which are easier to solve. Lastly, redistricting within the GIS environment in this research study will certainly ease the redistricting process with the ready-to-use functionality and storage capability.

3.3 Architecture Design of Proposed Shape-based Redistricting Model

The architectural design of the proposed model is shown in the semantic diagram in Figure 3.3. There are three main data stores for the proposed model, which provide the input data accordingly to the redistricting goals such as for election, school, forest or so on. Shape Optimal Rules (SOR) is the data store, which consists of the knowledge acquisition from the district planners or Decision Makers (DM). This data store includes the objective and the fuzzy rules sets in the FMCDM and another two data store: Application Dependent Data Store (ADD) and Application Independent Data Store (AID) that provide the other necessary input data for the proposed model.

The initiate stage of the proposed model is in Data Preparation Module (DPM), followed by three stages (S_1, S_2, and S_3) of the DP. At each of the stage, redistricting process on creating district occurs and decisions are made with FMCDM as the outcome of the redistricting process. Unqualified result will be sent to the following stage for the redistricting process meanwhile the qualified result will be sent to Combine Optimal District Module (COD). Finally, all districts are processed and combined again in the COD to get the final optimal district plan as mentioned in Section 3.1. The details of each component or module are discussed in the following section.

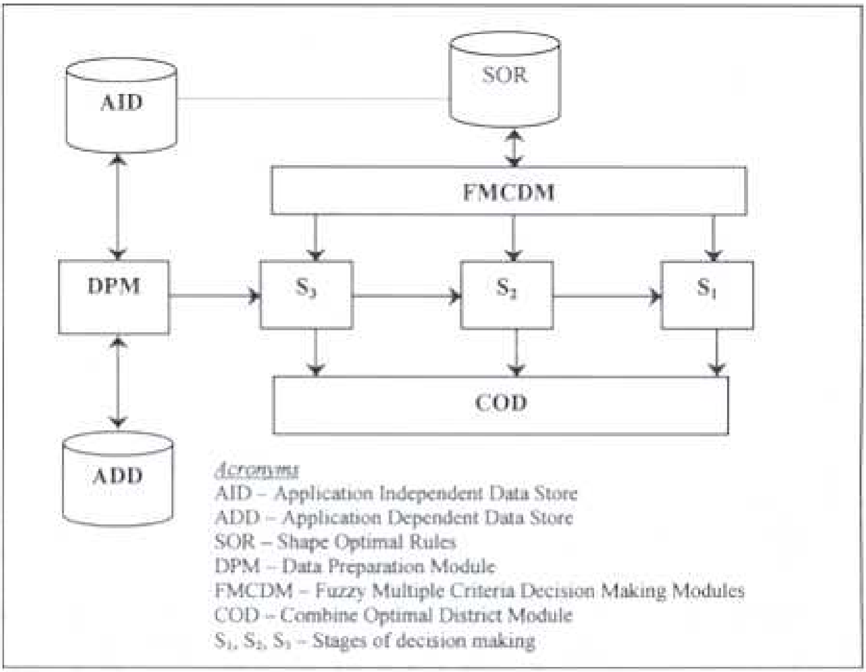


Figure 3.3: Semantic Diagram of the architectual design for the proposed model

3.3.1 Data Preparation for the Input

The first module of the proposed model is to prepare and group the spatial data into smallest features to represent the redistricting criteria in the attribute or non-spatial data (Figure 3.4) according the redistricting goals. For instance, the data input may be the population block in election boundary redistricting or police beat in Law Enforcement Redistricting. This input data format is in polygon and point features that are the fundamental representation techniques in geographic arena.

Preparation of all the spatial data will be needed for the first step of the model so that it can work with appropriate datasets is in polygon feature. Input data in polygon feature is suitable in redistricting process because polygons represent area features and they are made up of arcs, which define the boundary, and a label point, which links the polygon feature to an attribute record. GIS software like Arc\Info stores polygons topologically as a list of arcs and a label that makes up each polygon. Polygon topology here is important for the relationship manner (left, right or adjacent) between them. Besides, it is flexible and interchangeable to regions and points features. Furthermore, various kinds of geo-processing activities like polygon merging, intersecting, and overlaying can be conducted easily and efficiently with this polygon feature.

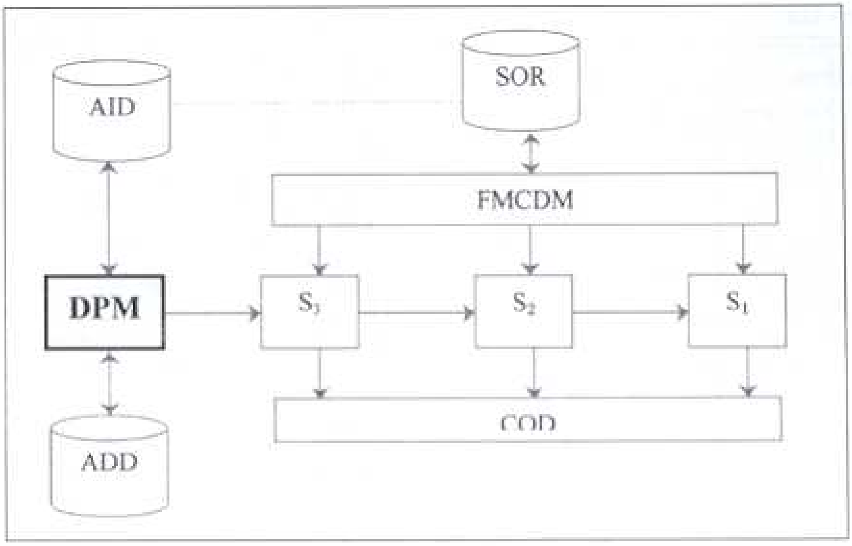


Figure 3.4: The initiate stage (Data Preparation Module) in the proposed model

This study determines to use triangles as the input polygon features. Triangles are made from three points, which can occur at any location. The reason to use triangles as the input polygon features is because triangulation is the mother of all polygon partitioning problems triangulation, and the interior of all kind of polygons can be completely partitioned into triangles (Skiena, 1997). In addition, the focus of the research is more on the decision support.

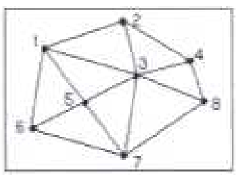


Figure 3.5: Triangles in triangulated irregular network data model

Consequently, after the preparation process, the polygon coverage will be ready with necessary attribute data depending on the application nature in order to be used in the following process. Instead of polygon, point feature are feasible as input data because it can be easily converted to polygons by Thiessen Polygons Technique (refer Figure 3.6) which is a ready-to-use technique in most of the GIS software like Arc\Info. Therefore, after the preprocessing process, the polygon coverage will be ready to be used. Some examples for the input data are like population distribution polygon or point for political redistricting, Police Reporting Areas (PRAs) polygon for law enforcement redistricting and student distribution polygon or point for school district. After the preprocessing process, the redistricting area will be intersected with a license boundary, which provides the exact location of the redistricting process.

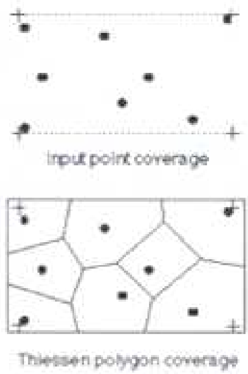


Figure 3.6: Input point to polygon

3.3.2 Shape Optimal Rules

Shape Optimal Rules (SOR in Figure 3.7) is the module that plays an important role to determine the optimality of district compactness. The SOR module will apply the concept of fuzzy set theory in IF-THEN ruling process (refer Appendix A.2 for details). The compactness measurement methods give a gray area (Figure 3.8) for district planners or decision makers on the district shape assessment and therefore fuzzy set theory will be used to formalize the gray area in order to specifically express them. In other words, fuzzy set theory is suitable and flexible to describe the vague value of this gray area. Furthermore, fuzzy numbers are intuitively easy to use in expressing the decision maker's, qualitative assessments (Deng, 1999; Chen, 1996; Mon, 1995; Yeh & Deng, 1997). Therefore, this module helps define the necessary shape optimal rules to be used later in the FMCDM module in order to obtain the Integrated Compactness Indexing. Then, this indexing will be used at each of the Dynamic Programming modules (S_1, S_2, and S3) in order to acquire the optimal compact districts.

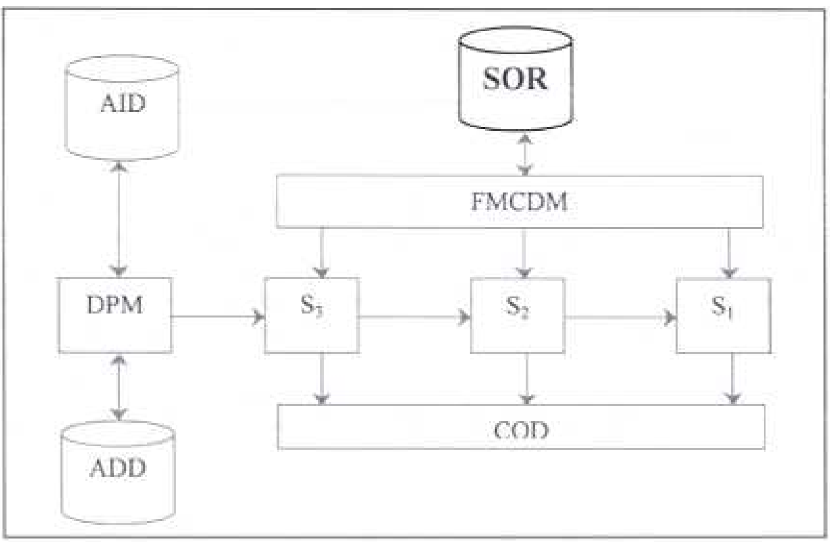


Figure 3.7: Shape Optimal Rules (SOP) for the proposed model is highlighted.

Figure 3.8: The Decision Making Process

The SOR component is related to the application independent data, which is the focus of this research on the geographical aspect as mentioned in requirement specifications and illustrated in the dot line in Figure 3.7. Therefore, this research uses compactness measurement method for generating the SOR. Although there are many compactness measurement methods as discussed in Section 2.4, this research chooses two methods to determine the SOR in order to measure the district boundary complexity and district shape compactness. Those two methods include one of the Euclidean Measure methods and one of the non-Euclidean Measure methods. Each of these compactness measurements represents a criterion in the multiple criteria decision making process later. The purpose for the integration of more than one compactness measurement methods is to combine the strengths of each measurement to the district compactness and to reduce their weaknesses by weighing vectors, which will be incorporated into the Integrated Compactness Indexing (ICI) in the FMCDM module later.

3.3.2.1 Selected Criteria for Shape Optimal Rules

Specifically, this research determines two different compactness measurement methods to consider the compactness criteria and they act as two different separate criteria. The two methods include the Euclidean Measure based on fractal dimension and the non-Euclidean Measure based on area-perimeter ratio. The following sections briefly explain the two selected criteria for the generation of shape optimal rules, specifically for shape-based redistricting.

3.3.2.1.1 Fractal Dimension

The higher the fractal dimension, the more complex it is for the district boundary. Subsequently, the district compactness is determined by the value on the fractal dimension. Thus, this research chooses to use box counting dimension to calculate the fractal dimension because of its versatility in the type of measurement (boundary, surface or lacunarity) and its ease of implementation. Furthermore, the box-counting dimension applies to any structure in the plane and can be adapted for structures in 3-dimensional space (Bowdren, 1997). Covering an image by 'r'-size boxes and determining how many boxes of a particular size 'r' are needed to cover the image that the box-counting dimension is determined. Thus, the number of boxes of size 'r' needed to cover the image is given by

$$N(r) = \frac{i}{r^D} \qquad (3.1)$$

To estimate the box-counting dimension, the Euclidean space containing the image can be divided into a grid of boxes of size 'r' and counting such boxes N(r) which are non-empty (Figure 3.9). Then, the size 'r' is changed to progressively smaller sizes and the corresponding numbers of non-empty boxes are counted N(r). The logarithm of N(r) versus the logarithm of 1/r gives a line whose gradient corresponds to the box dimension. The sequence of mesh sizes for grids is usually reduced by a factor of 1/2 from one grid to the next. Therefore, if the number of boxes counted increased by a factor of when the box size is halved, then the fractal dimension is equal to D (Bowdren, 1997).

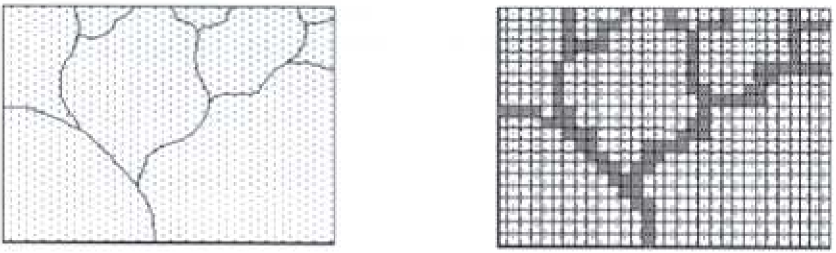


Figure 3.9: A grid of boxes of size 'r' and counting such boxes N(r).

Five different box sizes will be fixed and they are used to calculate the number of boxes that will intersect with the district boundary line. The box sizes and number of boxes intersected for selected district boundary are stored in temporary files and later converted to logarithm format. This data is then used to calculate fractal dimension in a graph. When the data that are plotted on a log-log graph (Figure 3.10) based on the results, will closely correspond to the following linear relationship:

$$\log(\text{box count}) = a + b \log(\text{box size}) \qquad (3.2)$$

where box count is the number of boxes overlapping the feature, and box size is the length of one side of the box. The intercept is represented by a, and b is the slope. The fractal dimension D is represented by the absolute value of the slope b. Subsequently, the slope b or in other word the fractal dimension D will be used to evaluate the district boundary curve.

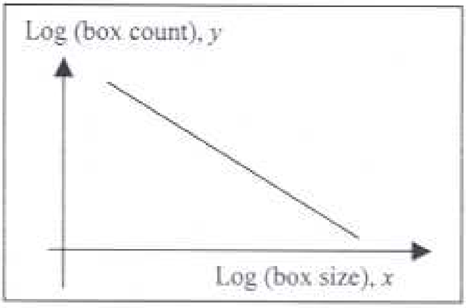


Figure 3.10: The linear relationship for the log-log graph

The value of fractal dimension was calculating from the slope from the formula below which is adopted from Microsoft Excel to calculate the slope.

$$\textbf{Fractal Dimension = slope of the log-log graph} = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2} \qquad (3.\ 3)$$

3.3.2.1.2 Euclidean Measure based on Area-perimeter ratio

Another Euclidean compactness measurement used to control the inconsistency of the fractal dimension is compactness measurement based on Area-perimeter ratio. It is efficient to consider fractal dimension in measuring the district boundary complexity and to incorporate the stability of this Euclidean Measure based on area and perimeter. The Euclidean Measure for the compactness is as below:

$$compactness = \frac{Perimeter^2}{Area} \qquad (3.\ 4)$$

This method is stable and performs very consistent with the changes of the district shape. A preliminary experiment is conducted (refer Section 4.2.1) and the result proves the consistency of this method comparing to the fractal dimension in the non-Euclidean Measure. Both of these measurements are used together in the case study to provide the new Integrated Compactness Indexing (ICI) value for the use in the FMCDM in the next section.

3.3.2.2 Definition for Shape Optimal Rules

After selection of the criteria for the shape optimal rules, this research defines the rules by linguistic terms and membership function. Linguistics terms are defined for each of the selected criteria and their weighing vector. The weighing vector represents the decisions of the district planners at the importance on each criterion among all the criteria. The fuzzy set theory will be applied in these weighing vectors for the ranking of the relative importance's (which is another portion of gray area) during the decision making process. This research also uses the definition of the interval of confidence at level α (α-cuts) the helps characterize the triangular fuzzy number as

$$\begin{aligned} &\forall \alpha \in [0,1], \\ &\overline{A}_n = [a_1^n, a_3^n] = [(a_2 - a_1)\alpha + a_1, -(a_3 - a_2)\alpha + a_3] \end{aligned} \qquad (3.\ 5)$$

where $\overline{A}$ is a triangular fuzzy number with $\overline{A} = (a_1, a_2, a_3)$ where a_2 is the most possible value of fuzzy number A, and a_1 and a_3 are the lower and upper bounds respectively which is often used to illustrate the fuzziness of the data evaluated (refer Appendix A.2).

However, districts that made up by the natural boundary like island, river, peak of a mountain and others will be defined as compact district boundary line because of their unchangeable nature. In other words, this research defines district boundary line as compact although a district is non-contiguous or questionably contiguous due to the natural boundary (Section 1.2). Therefore, the definition in this section does not consider those boundaries those are made up with the unchangeable natural boundary.

Subsequently, effectiveness of the selected compactness measurement methods will be obtained from DM. However, DM may often meet the situation where it is difficult for them to choose or reject alternatives. Thus, only the YES/NO method needs to be improved due to the subjective judgments of the DM (Ron, 1998). One alternative method is for the DM to give an effectiveness level (x) to each of the selected criteria (C). The effectiveness levels belong to a set of linguistic terms that contains various degrees of preference required by the DM. Linguistic terms are words or sentences in natural or artificial languages such as "very low", "low", "medium", "high", "very high" for comparison of human height. Linguistic terms are ill defined and can hardly be described by single numerical values. Then, this research able to model the subjectiveness and vagueness of the decision making process by linguistic terms which is found intuitively easy to use according to Deng (1997). This research will use linguistic terms μ (effectiveness, x) = {Very Poor (VP), Poor (P), Fair (F), Good (G), Very Good (G)}, defined in Figure 3.11. A membership function, which assigns to each linguistic term with a grade of membership, is associated with the fuzzy set in [0, 1]. When the grade of membership for a linguistic term is one, it means that the linguistic term is absolutely in that set. When the grade of membership is zero, it means that the linguistic term is not in that set. Borderline cases are assigned values between zero and one.

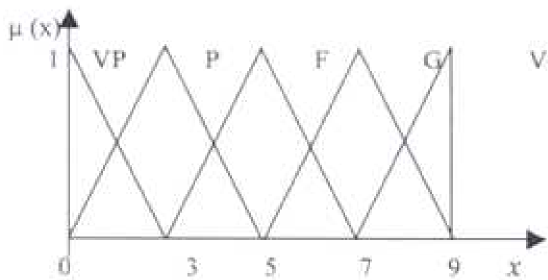


Figure 3.11: Membership functions, μ(x) and linguitics terms for effectiveness (x) of each criteria (C)

Considering the different importance of criteria, different weights are determined for each criterion. By pairwise comparison of the relative importance of criteria, the pairwise comparison matrix $E = [e_{ij}]_{t\times t}$ is established, where e_{ij} represents the quantified judgments on pairs of criteria C_i, and C_j. The comparison scale ranges from 1 to 9, each representing the concepts of: 1: equally important; 3: weakly more important; 5: strongly more important; 7: demonstratively more important; 9: more important; 2, 4, 6, and 8 are intermediate values between adjacent judgments. For example, $e_{ij} = 5$ means C_i is strongly more important than C_j. To facilitate the making of pairwise comparison, triangular fuzzy numbers defined in Table 3.1 are used. A triangular fuzzy number $\bar{x}$ expresses the meaning of 'about x', where $1 \le x \le 9$, with its membership function. Fuzzy number $\bar{x}$ used by Deng (1999) is revised here to better reflect the decision situation involved.

Table 3.1: Linguistic terms used by the decision matrix

Linguistic term	Very Poor (VP)	Poor (P)	Fair (F)	Good (G)	Very Good (VG)
Fuzzy number	$\bar{1}$		$\bar{x}$		$\bar{9}$
Membership function	(1, 1, 3)	$(x-2, x, x+2)$ *for*		$x = 3,5,7$	(7, 9, 9)

3.3.3 FMCDM by using Fuzzy-AHP

This section provides a metaphor to integrate multiple compactness measurements with FMCDM method to produce an Integrated Compactness Indexing (ICI) as the decision indexing for the DM according to the fuzzy rulesets from the SOR in previous section. Two compactness measurements are taken into account for the criteria and each of them is assigned with a weighing vector to determine their importance towards the overall redistricting process. The integration process is shown in the highlighted module in Figure 3.12.

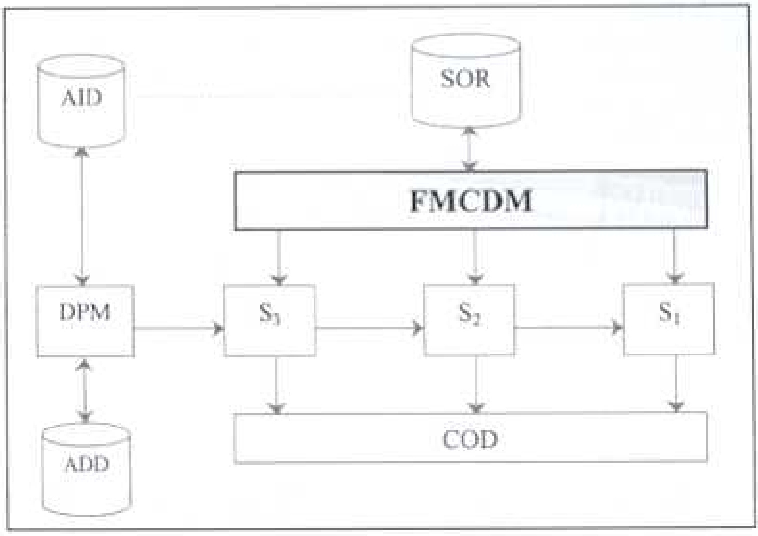


Figure 3.12: Highlighted portion is for the Fuzzy Multiple Criteria Decision Making Modules.

This research chooses to use a FMCDM method that called Fuzzy Analytical Hierarchy Process or Fuzzy AHP for the integration. Analytical Hierarchy Process is a multiple criteria method which uses hierarchic structures to represent a decision problem and then develops priorities for the factors based on decision maker's judgement. It is widely used to solve complicated, unstructured decision problem and thus it should be concerned with the processing of fuzzy information (Deng, 1999). It is difficult to get exact ratios for a pair of criteria considered, thus fuzzy ratios for the relative significant may incorporates the natural feelings of human beings. In other words, fuzzy theory is effective when the situation contains fuzziness from human subjectivity in redistricting functions. Indeed, Fuzzy AHP method was discussed and being used for ranking of Indian Coals in industrial use (Ravi and Reddy, 1999), rate and ranking the disability (Jacob and Zhesheng, 1997), assessing risk of cumulative trauma disorders (Pamela and Heng, 1997), evaluating weapon system (Mon, 1995) and many other applications. Consequently, this method is a useful method in enhancing the compactness indexing for shape-based redistricting because the AHP allows DM to express their judgements of pairwise comparison in fuzzy ratios for indicating its importance for each criterion. In addition, the fuzzy ratios is able to avoid unbalanced scale of estimations and its ability to adequately handle the uncertainty and imprecision associated with the mapping of the decision makers' perception to a crisp number (Deng, 1999). Consequently, this research will solve the shape-based redistricting problem by the use of the ICI in Section 2.4.3 and achieve the requirement specifications in Section 3.1. Fuzzy-

AHP not only helps consider multiple criteria when solving the shape-based redistricting problem but also helps rank the compactness of the district plan by the ICI.

This model adopts the Fuzzy AHP method from Ravi and Reddy (1999), Jacob and Zhesheng (1997), Mon (1995), Pamela and Heng (1997) for the proposed model. Therefore, the shape-based redistricting decision problem consists of (a) a number of alternatives, which refer to the individual district, denoted as A_i $(i = 1, 2, \ldots, n)$, (b) a set of evaluation criteria C_j $(j = 1, 2, \ldots, m)$ refers to the compactness measurements and other application dependent data, (c) a qualitative assessment x_{ij} $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m)$ (referred to as *performance ratings*) representing the performance of each alternative *Ai* with respect to each criterion C_j, leading to the determination of a decision matrix for the alternatives, and (d) a weighting vector (referred to as *criteria weights*) representing the relative importance of the evaluation criteria with respect to the overall objective of the problem.

Then, this research uses a FMCDM approach based on the synthesis of the following concepts, including (a) fuzzy set theory, (b) AHP, (c) α-cut concept on the level of confidence, and (d) Decision maker (District Planner), DM's attitude towards risk. Consequently, the subjectivity and imprecision of the evaluation process are adequately handled, and the complex and unreliable process of the ranking procedure starts at the determination of the criteria importance and alternative performance. By using the fuzzy numbers defined in Table 3.1, a fuzzy reciprocal judgement matrix for the decision matrix for m criteria and n alternatives is given as

$$X = \begin{bmatrix} x_{11} & x_{12} & \ldots x_{1m} \\ x_{12} & x_{22} & \ldots x_{2m} \\ \ldots & \ldots & \ldots \ldots \\ x_{n1} & x_{n2} & \ldots x_{nm} \end{bmatrix} \qquad (3.6)$$

where x_{ij} represent the linguistic assessments of the performance rating of alternative $A_i (i = 1,2,\ldots n)$ with respect to criterion $C_j (j = 1,2,\ldots m)$. The decision matrix is to be given by the DM based on the term set being defined in Figure 3.11.

The weighting vectors for the evaluation criteria can be given directly by the DM or obtained by using pairwise comparison of the AHP. Linguistic terms to express the weighing vectors are defined in Table 3.1. The weighting vectors W for the criteria is as follow:

$$W = (w_1, w_2, \ldots, w_m) \qquad (3.7)$$

Then, the model obtains a fuzzy performance matrix (3.8) representing the overall performance of all alternatives, with respect to each criterion can by multiplying the weighting vector by the decision matrix as $z = w \times x$. The arithmetic operations on these fuzzy numbers are based on interval arithmetic.

$$Z = \begin{bmatrix} w_1 x_{11} + w_2 x_{12} + \ldots + w_m x_{1m} \\ w_1 x_{12} + w_2 x_{22} + \ldots + w_m x_{2m} \\ \ldots \quad \ldots \quad \ldots \quad \ldots \\ w_1 x_{n1} + w_2 x_{n2} + \ldots + w_m x_{nm} \end{bmatrix} \qquad (3.8)$$

By using a α-cut (3.5) on the performance matrix (3.8), an interval performance matrix can be derived as in (3.9), where $0 \leq \alpha \leq 1$. The value of α represents the DM's degree of confidence in his/her fuzzy assessments regarding alternative ratings and criteria weights. A larger α value indicates a more confident DM, meaning that the DM's assessments are closer to the most possible value a_2 of the triangular fuzzy numbers (a_1, a_2, a_3).

$$Z = \begin{bmatrix} [z_{1l}^{\alpha}, z_{1r}^{\alpha}] \\ [z_{2l}^{\alpha}, z_{2r}^{\alpha}] \\ \cdots \\ [z_{nl}^{\alpha}, z_{nr}^{\alpha}] \end{bmatrix} \qquad (3.9)$$

Incorporated with the DM's attitude towards risk using an optimism indexing λ, an overall crisp performance matrix is calculated as in (3.10), where $z_{i\alpha}^{\lambda'} = \lambda z_{ir}^{\alpha} + (1-\lambda) z_{il}^{\alpha}, \lambda \in [0,1]$

$$Z_{\alpha}^{\lambda'} = \begin{bmatrix} z_{1\alpha}^{\lambda'} \\ z_{2\alpha}^{\lambda'} \\ \cdots \\ z_{n\alpha}^{\lambda'} \end{bmatrix} \qquad (3.10)$$

In this proposed shape-based redistricting model, $\lambda = 1$, $\lambda = 0.5$, and $\lambda = 0$ are used to indicate that the DM involved has an optimistic, moderate, or pessimistic view respectively. An optimistic DM is apt to prefer higher values of his/her fuzzy assessments, while a pessimistic DM tends to favor lower values.

The evaluation criteria or compactness measurement indexing in this model is a relative measure and not an absolute measure. Therefore, the evaluation of the similarity of relative compactness judgments between each pair of measures will consider a normalization process. Therefore, after the facilitation of the vector matching process, a normalization process in regard to each criterion is applied to (3.10) by using (3.11), resulting in a normalized performance matrix expressed as in (3.12).

$$z_{i\alpha}^{\lambda} = \frac{z_{i\alpha}^{\lambda'}}{\sqrt{\sum_{i=1}^{n} (z_{i\alpha}^{\lambda'})^2}} \qquad (3.11)$$

$$N_{\alpha}^{\lambda} = \begin{bmatrix} z_{1\alpha}^{\lambda} \\ z_{2\alpha}^{\lambda} \\ \cdots \\ z_{n\alpha}^{\lambda} \end{bmatrix} \qquad (3.12)$$

The values of N_{α}^{λ} indicate the degree of preference with respect to the alternatives for fixed α and λ, respectively where $\alpha \in [0,1], \lambda \in [0,1]$. Indeed, this value is the Integrated Compactness Indexing (ICI), which considers all evaluation criteria earlier. Therefore, the larger the value, the more the preference of the alternative.

In summarizing the Fuzzy-AHP method for getting the Integrated Compactness Indexing above, this research presents the steps required for the algorithm in Figure 3.13.

(1) Formulate redistricting decision problem as multiple criteria problem
(2) Identify the hierarchical structure of the problem
(3) Define membership function for each criteria
(4) Apply IF-THEN decision rules to assign a linguistic and numeric value of each criteria variable
(5) Obtain decision matrix by fuzzy number as expressed in (3.6) using AHP method based on fuzzy number defined in Table 3.1 and Figure 3.11
(6) Obtain weighting vector for the criteria as expressed in (3.7) using AHP method based on fuzzy number defined in Table 3.1 and Figure 3.11
(7) Obtain fuzzy performance matrix (3.8) by multiplying the decision matrix obtained at step 5 by the weighting vector determined at step 6
(8) Obtain interval performace matrix (3.9) by α-cuts on the performance matrix determined at step 7
(9) Obtain crisp performace matix (3.10) by DM's attitude towards risk represented by an optimism indexing, λ
(10) Calculate normalized performace matrix (3.12) by (3.11)
(11) Get the Integrated Compactness Indexing (ICI)

Figure 3.13: Algorithm for the Fuzzy-AHP for this study

3.3.4 Dynamic Programming (DP)- Problem Solving by Stages

Later, the problem will be divided into sub-problems of the same kind called stages. These stages will be the logical sub-problems, which consists of input variables or state variables those are the possible beginning situations or conditions of a stage. The next step (at Stage 3, S_3) is to solve the last or the worst stage of the problem for all possible conditions or states using an array or sequences of arrays. Solving the last or worst stage is trivial and it acts as the preprocessing period to provide data to the decision scheme in FMCDM. Then, the scheme working backward from the last stage and solve each intemediate stage (at Stage 2, S_2 and Stage 1, S_1). This is done by the optimal policies determined in FMCDM from that stage to the end of the problem. The decision to make at every stage aims to reach the optimal compact (Refer Appendix A.3 for detail). Therefore, the decision variables or decision indexing, which is ICI, that exist at each stage are regarding to the district geometry shape.

The dynamic programming used in this research is briefly discussed in Section 3.3. It is a problem solving method by using stages or called sub problems. There is overlapping of data used in different sub-problems and these overlapping data is the advantage to avoid NP-complete situation. This method will solve the last or worst stage first. Although the first solution or recursive is trivial but it preprocesses and provides data to the second stage where the optimality will begin by the optimal rules as mentioned earlier.

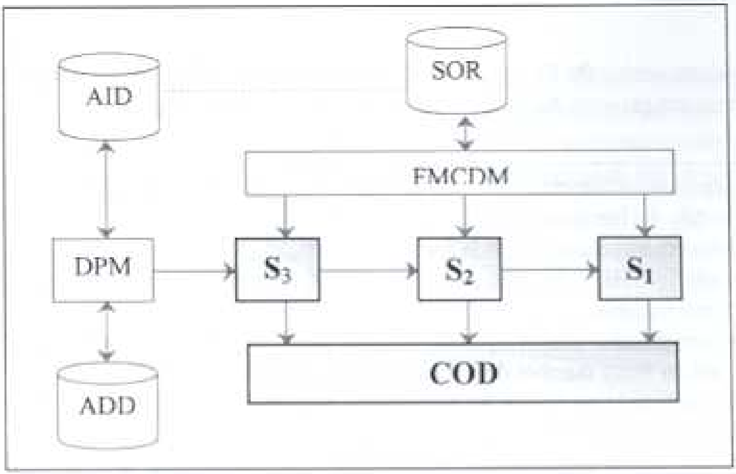


Figure 3.14 Highlighted portion as the Dynamic Programic Modules

After the data is prepared from the previous module (see Figure 3.14), the original problem will be divided into sub-problems called stages. The dividing of stages will depend on the distribution of the data input variation by finding the maximum, minimum and mean of the redistricting criteria. Secondly, the problem will be solved from the last or the worst stage of the problem for all possible conditions or states. Then, the process continues with working backward from the last stage. When solving each intermediate stage, decision will be made from that stage to the end of the problem. The decision making is made through checking on the relative district shape by the Integrated Compactness Indexing (ICI) by the FMCDM module (refer Section 3.3.3). Finally, the process ends with obtaining the optimal solution for the original problem by solving all stages in the following module. They are to help to set up and solve the redistricting problem. These equations are then used in Figure 3.15 to clarify each stage for dynamic programming. The mathematics notation to describe the important concepts of the dynamic programming is as following equation.

S_n = **input to stage** n **(3. 13)**

D_n = **decision at stage** n **(3. 14)**

R_n = **return at stage** n **(3. 15)**

S_{n-1} = **output from stage** n **(3. 16)**

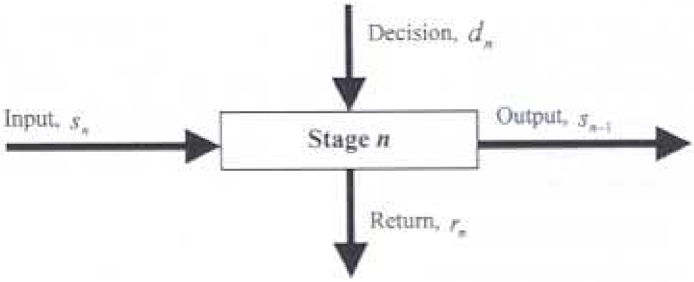


Figure 3.15: Input, Decision, Output, and Return for DP Schema (Render, 1997)

Transformation for dynamic programming in this research is the location/allocation process performs the redistricting process through annexes or merges adjacent triangular to create district after the proposed model sorts and selects the polygon at each stage. The process continues with checking and merging the smallest polygons until it reaches a polygon with the ideal area size. This technique is adapted from Automated Redistricting System for Law Enforcement in Maryland (Fred, 1997). This transformation process allows the model to proceed from one stage to another. The transformation function for stage 2 (t_2), convert the input to stage 2 (s_2), and the decision made at stage 2 (d_2) to the output from stage 2 (s_1). It can be represent as t_2 (s_1, d_2) because the transformation process depends on the input and decision at any stage. In general the transformation function can be represented as in equation (3.17). The general formula allows the solution to proced from one stage to another using te transformation function in (3.18). Another useful quantity is the total return at any stage. The total return allows us to keep frame of the total district at each stage as the research solves the dynamic programming. It can given as in equation (3.19).

$$t_n = \text{transformation function at stage } n \qquad (3.17)$$

$$s_{n-1} = t_n\ (s_n\ ,\ d_n) \qquad (3.18)$$

$$f_n = \text{total return at stage } n \qquad (3.19)$$

Then, every return, f_n from the stage n will be sent to Combine Optimal District (COD) module. Thus, the COD module acts to receive the return from every single stage and the final output of the decision stages in order to combine the result together to get a complete and optimal district plan. Existing boundary like licensed boundary and rivers are important to be taken into account as one of the considerations of the proposed model (Section 3.1). Therefore, these layers will be overlaid onto the polygon layers in order to create the district boundary line. This process may not influence too much to the district shape compactness and continuity because the research defines the district boundary line, which made up by natural boundary as compact district. Therefore, the output will reach to optimal condition where the entire optimal compact districts from each individual stage are gathered together.

3.4 Summary

The research aims to integrate Fuzzy Multiple Criteria Decision Making in a new compactness measurement indexing and apply it with Dynamic Programming for the shape-based redistricting within GIS environment. This is due to the requirement specifications that aim to integrate multiple compactness measurement method to produce an improved and integrated shape assessment indexing, which copes with fuzziness. Besides, the new indexing will be incorporated into redistricting process or algorithm to generate an optimal compact district within an GIS environment that is able to manage the spatial aspect in redistricting. The architectural design of this model is easy to understand with the dynamic programming approach, which divides the sub-problems into smaller instances. Multiple criteria consider human subjectivity in a more systematic manner with the help of fuzzy concept. In addition, the result of the district plan too can be ranked under various degrees of confidence in assessment with different attitudes toward risks. Consequently, the DM can better understand the problem itself and the implication of decision and subsequently improve their confidence in making decisions. The integration of the multiple criteria Fuzzy Decision Making process is believed will helps to decide the optimal alternatives among redistricting plans from the computer-generated redistricting algorithm.

CHAPTER 4: INTEGRATED COMPACTNESS INDEXING AND PROTOTYPING

After the requirement specifications and the conceptual design have been identified from the previous chapter, this chapter discusses the development of the Integrated Compactness Indexing (ICI) and a prototyping of the proposed shape-based redistricting model. The proposed model will integrate multiple compactness measurement method to produce the integrated shape assessment indexing namely ICI that is able to cope with fuzziness. The new indexing will be incorporated into redistricting process or algorithm to generate an optimal compact district within an GIS environment that is able to manage the spatial aspect in redistricting. Therefore, the proposed model will integrate the concept of Fuzzy-AHP with Dynamic Programming for a shape-based redistricting process within GIS environment.

In brief, the prototype development workflow is illustrated in Figure 4.1. The development process consists of steps like prototyping setup, knowledge acquisition, prototype development, and testing and evaluation. At first, this research conducts a prototyping setup in order to identify a specific case study and a required platform for implementation. The setup would produce a testbed for the proposed model according to the conceptual design in the previous chapter. After that, during the knowledge acquisition process, knowledge is originated from several reports and case studies from previous researches directly and indirectly regarding to the shape-based redistricting research analysis and with a preliminary survey. This knowledge is then used for the integration of Fuzzy-AHP to generate the ICI for the development of the enhanced shape-based redistricting with Dynamic Programming technique. Later, the design and development of the prototyping are elaborated in form of data flow diagrams to illustrate the systematic prototyping process. Finally, testing and evaluations are conducted to handle errors and to improve the overall performance for the prototype.

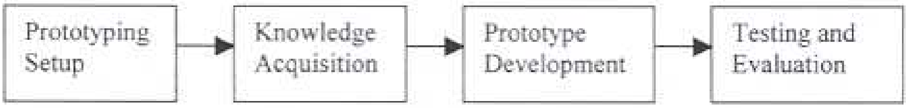


Figure 4.1: Prototype Development Workflow

4.1 Prototyping Setup

The setup of the prototyping in the implementation process refers to the necessary requirement specifications for the prototyping. Therefore, a case study would be selected in order to verify the practically of the enhanced model. The selected case study for implementation is on Forestland Blocking for dividing the forestland into different blocks according to specific criteria for the purpose of enrichment planting, industry planting and so on. After that, the required platform is determined in order to have a suitable hardware and software interface.

4.1.1 Selected Case Study

A specific case study on Forestland Blocking is considered to provide requirements for all the aspects of the enhanced model and provide a test-bed for implementation. The case study for this research is extremely important in forestry to ensure sound control and allocation of expenses. It is an operation in planning to divide forestland from a large area into smaller one so that the activities like land clearing, planting or maintenance pertaining to a forest operation is accounted to the smaller area. The main factors to the Forestland Blocking are slope and known features like river and license boundary. All these factors are important to maintain the compactness and the continuity of the smaller area later. In common, an area of about 80,000 hectare is divided into compartment with around by slope and later into blocks by existing boundaries.

Therefore, the input data for the prototype are including records on the elevation for contour because it will be used as the application dependent factors to draw the district line. Projection information is assumed as complete and all geometry data such as area and perimeter for each polygon is considered accurate.

4.1.2 Selected Hardware and Software Environment

In brief, the physical environment of the prototype is within GIS software called ARC\INFO 7.2.1 from ESRI, which will provide most of the spatial function needed for the prototype development process. Then, the basic hardware environment included Window NT 4.0 with 32 MB RAM and Hard Disk 2.0 GB. However, a more sophisticated hardware environment will certainly fasten the process in the model itself. The language used to implement the model is Arc Macro Language (AML), which is the scripting language, used in the ARC\INFO. The extensions in ARC\INFO that must be used include ARCEDIT, ARCPLOT, TIN, GRID and TABLES.

4.2 Knowledge Acquisition for Integrated Compactness Indexing (ICI)

One of the advantages of integrating Fuzzy-AHP to shape-based redistricting is the capability to assist the Decision Maker (DM) in prioritizing or selecting one or more alternatives from a finite set of available ones with respect to multiple, usually conflicting criteria. With the characteristics of the redistricting problem, Fuzzy-AHP is well suited for evaluating the overall performance of a specific shape-based redistricting task. However, selecting the suitable criteria depends on the redistricting goals at the very beginning stage. For instances, when Fuzzy-AHP is used for election boundary redistricting, the considered criteria are the population, race, income, compactness, and continuity. It is the same for other applications like school redistricting, law enforcement redistricting and so on. Therefore, the following subsections discuss the details regarding the knowledge acquisition process in the prototyping. There is a preliminary experiment on the selected criteria in order to obtain knowledge of the ICI. After that, the research is able to define the shape optimal rules, which will help to optimize the overall district plan compactness in the redistricting process later.

4.2.1 Preliminary Experiment

An experiment is conducted in order to find the shape optimal rulesets of the selected criteria on the compactness. Thus, the experiment is conducted toward the criteria based on fractal dimension and based on area-perimeter ratio. This experiment is extremely important to obtain the interval value in x-axis of the membership function or the effectiveness level, x for each linguistic term as mentioned in Section 3.3.2.1.2. Therefore, this experiment will help conclude the exact value for definition of the linguistics term in the next subsections.

A set of 29 special polygons sample shown in Figure 4.2 is created to survey both of the compactness measurement indexing produced by the fractal dimension and the Euclidean Measure by area-perimeter ratio. The most compact district is the circle shape. Thus, detail information about the district compactness will refer to this most compact shape in order to describe the district shape. Some of the generated polygons have basic shape like rectangle, oval, triangle, and other nice shape polygon. Besides, this experiment creates some other odd-polygons that are having many angles, island, and so on. Different kinds of polygon shape are used in this experiment in order to get the performance of the selected compactness measurement on different shape of the polygon. In other words, the experiment result helps to gather the information regarding the most compact district, compact district, not compact district, and so on for the definition of the linguistics term of the shape optimal rules. Besides, the gray area for decision making (refer Section 3.3.2) on the optimal district would be analyzed properly and this will certainly help the DM on the vague value when analyzing the district shape.

This research considers two criteria regarding the compactness (discussed in Section 3.3.2.1), and other application dependent criteria like district size, and slope. The two criteria related to compactness consist of the non-Euclidean compactness measurement based on Fractal Dimension and the Euclidean compactness measurement based on area-perimeter ratio. In contrast to the latter method, it is difficult to use and implement fractal dimension in any GIS environment. However, a special package called GRID in ARC\INFO helps to provide a faster and easier way for calculating the fractal dimension from Box-counting algorithms, which will cut district coverage into smallest cells. Furthermore, Knight (1997) had conducted and proved the usability of GRID in ARC\INFO from the box counting algorithm in order to obtain the Fractal Dimension.

Figure 4.2: Sample set of district shape to be used to determine membership function.

This research starts the experiment by calculating the compactness indexing from fractal dimension and area-perimeter ratio. The calculation is done automatically through the programming logic and flow in Arc Macro Language in ARC/INFO. Therefore, this research creates a program in this GIS environment and it helps to generate the compactness indexing for both of the compactness measurement quickly. Some of the important flow and code is explained in this section to give a clearer view for this experiment. Calculation of the fractal dimension is conducted by using GRID in ARC\INFO. Besides, some other subcommands are used to create a grid from polygons and lines in ARC/INFO coverage (Figure 4.3).

```
POLYGRID <in_cover> <out_grid> {value_item} {lookup_table} {weight_table}
LINEGRID <in_cover> <out_grid> {value_item} {lookup_table} {weight_table}
```

Figure 4.3: Command to convert the line and polygon grid by desired cell size.

```
SELECTMASK(<grid>, <mask_grid>)
```

Figure 4.4: Command to create boundary grid by a mask

Then, SELECTMASK command is used to create the boundary by masking all cell locations in the first input grid that have been assigned NODATA in the second input grid on a cell-by-cell basis within the analysis window (Figure 4.4). In the final grid file, a Value Attribute Table (VAT) is important for storing attributes defined on the set of values in the grid. Each unique value in the grid corresponds to one record in the VAT. The VAT always contains two integer items, VALUE and COUNT. Therefore, the VALUE item is the district internal number in ARC\INFO meanwhile the COUNT item represent the total of cells for a single district in each VALUE. By changing the cell size, there will be five different set of value for creating the linear log-log graph after the values converted to Logarithm forms (Section 3.3.2.1). After that, the fractal dimension is obtained from the slope of the linear graph according to the equation (3.3).

Table 4.1: Result for two different measure for district shape in Figure 4.2

District Number	*Euclidean Measure*	*Fractal Dimension*
2	43.3004	1.0126
3	14.0527	1.0102
4	12.5779	0.9841
5	16.0017	1.1261
6	45.6757	0.9893
7	40.9493	1.0043
8	16.5371	1.0089
9	19.9991	1.0562
10	20.5611	1.0155
11	30.7006	1.0024
12	72.6556	1.0127
13	40.8622	0.9918
14	19.7707	1.0084
15	16.0162	0.9949
16	113.3842	1.0278
17	23.6295	1.0102
18	39.8167	1.0070
19	28.0176	0.9716
20	15.5493	0.9913
21	19.4320	0.8957
22	55.2553	0.9977
23	18.6523	1.0265
24	21.1447	1.0279
25	37.9112	1.0515
26	21.5037	1.0800
27	18.9608	1.0048
28	43.7079	1.0126
29	18.8888	1.0312

On the other hand, the formula to find Euclidean Measure is based on the equation (3.4) in Section 3.3.2.1 and the details of the coding are in Appendix A.3. The experiment's result is done in Table 4.1. The result indicates fractal dimension does not provide consistent result for the relevant polygon shape. For instance, polygon (6) is not compact but the fractal dimension value shows value of 0.9893, which is close to the value for the circle (value for polygon (4) is 0.9841). Polygon (3) and polygon (5) are more compact because they have value 1.0102 and 1.1261 respectively. Besides, the experiment shows that this measurement does not show extinct result for polygon, which has island like polygon (22). In other words, this measurement treats polygon with island as compact and it does not consider continuity factor. However, some polygons give quite well result in providing the district compactness and show gradually changes in the value. For examples, polygon (15) and polygon (20) are with 0.9949 and 0.9913 each. The reason for that inconsistency condition is due to the box count method used.

Meanwhile, the Euclidean Measure by area and perimeter performs well and consistent. The cases mentioned earlier for polygon (3) and polygon (5) show gradually changes based on the shape compactness. In addition, polygon (22) shows that Euclidean Measure based on area-perimeter ratio treat polygon with island as not compact. The preliminary experiment is successfully conducted and it shows the performance of each criteria related to compactness. After that, the result will help the determination of the shape optimal rules so that they could be used accordingly in the implementation stage later.

The knowledge acquisition in this preliminary experiment is important in this study for generation of the ICI to improve the existing compactness measures by integrating multiple compactness measures. With respect of ICI, this research takes into account for the strengths of particular compactness measure and reduces the weakness of the compactness measures by controlling the fuzzy weighting vector for each criterion, which related to the compactness measurement. Besides, this research incorporates the level of confidence and attitude of risk for the DM. These two parameters are the contribution of DM relative weighing vectors among all criteria.

4.2.2 Shape Optimal Rules

Application dependent factor is important to draw a district plan, and thus the prototyping for this research will not ignore the criteria for the forestland blocking like the slope and block size. However, due to the inconsistent distribution of slope value, it may certainly cause the odd district shape especially at the higher slope range. Therefore, this prototyping would only consider the block or district area size to produce the Integrated Compactness Indexing (ICI) together with the two selected compactness measurement as another two criteria. This research represents and clarifies this shape-based redistricting problem definition with a hierarchy structure as in Figure 3.2 to represent the goals, considered criteria and the output (Section 3.3.2.1). In simple, the evaluation during the shape-based redistricting process for the case study is summarized in the Figure 4.5 and it is relying on the district area size, fractal dimension that represents the district boundary complexity and the Euclidean Measure to represent the district compactness and continuity.

Therefore, the shape optimal rules are defined according to the result from the preliminary experiment and the concept in Section 3.3.2. Let assume the ideal area is A and this is the value determined by the DM, so the district with area A is defined and categorized as "Very Good (VG)" district. Then, the allowed range of district area is within B meter square. Therefore, the research defines as "Good (G)", "Fair (F)", "Poor (P)" and "Very Poor (VP)" for every quarter changes of B size. It can be illustrated more clearly in Figure 4.6. Besides, other compactness measurements for fractal dimension and Euclidean Measure based on area and perimeter is defined based on the experiment result as described in previous section. According to the experiment result, the most compact indexing for Euclidean Measure based on area-perimeter ratio is above 26. Meanwhile, the most compact indexing for fractal dimension is around 1.0. Therefore, the x-axis in the membership function is dependent on the most compact value and each of the linguistics term mentioned is the graduate change in the compactness indexing. Figure 4.7 shows the definition of district boundary complexity according to the FD criteria as {"Good (G)", "Quite Good (QG)", "Average (A)", "Quite Complex (QC)" and "Very Complex. (VC)"}. Meanwhile, the linguistic terms for district compactness indexing for EM is defined as {"Very Very Compact (VVC)", "Very Compact (VC)", "Compact (C)", "Not Very Compact (NVC)", "Not Compact (NC)"} in Figure 4.8. The shape optimal rules in Figure 4.6, Figure 4.7 and Figure 4.8 are needed in the FMCDM module in Section 4.3.3 later.

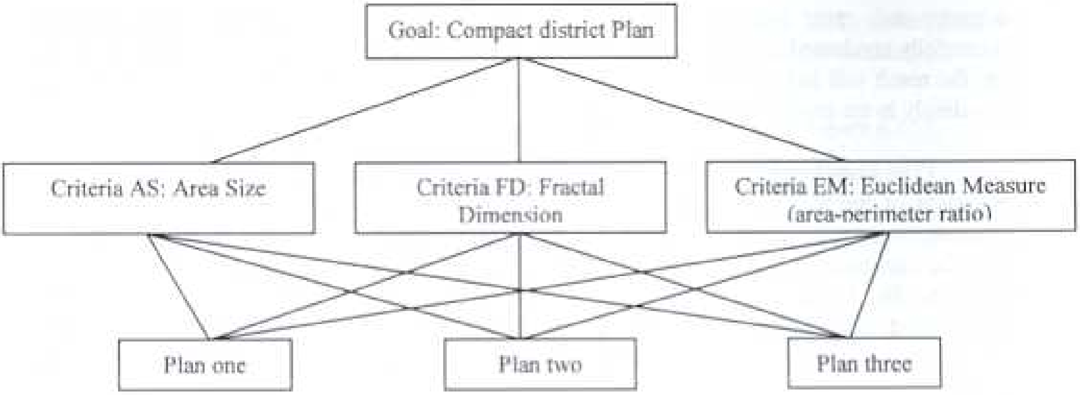


Figure 4.5: Three criteria are considered in the problem definition for forestland blocking

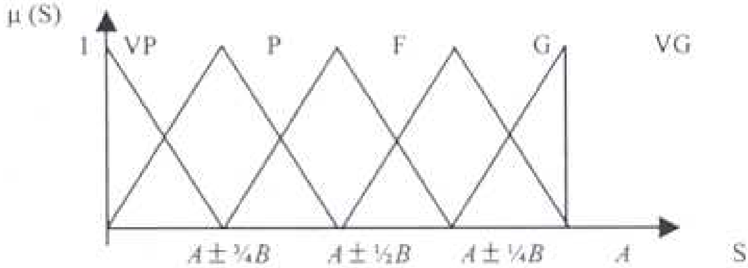


Figure 4.6: Membership functions, μ(S) and the linguitics terms for district area size, S

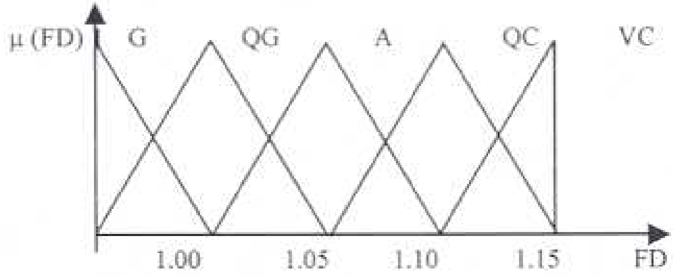


Figure 4.7: Membership functions, μ(FD) and the linguitics terms for Fractal Dimension, FD

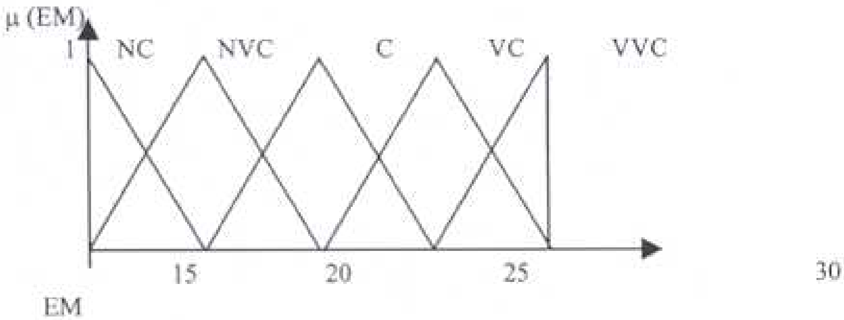


Figure 4.8: Membership functions, μ(EM) and the linguitics terms for Euclidean Measures, EM

4.3 Prototype Design and Development

The main stages of the prototype development in this research has four main modules as illustrated with Module One to Four in the Diagram 0 of the Data Flow Diagram in Figure 4.9. These modules are specifically named as Initialization module, Districts Conquering module with Dynamic Programming technique, Fuzzy Multiple Criteria Decision Making (FMCDM) module and Combine Optimal District (COD) module. Each of the modules has their specific functionality and contribution to the prototype. For instance, the Module One or Initiation module is to prepare the input data meanwhile the Module Two or District Conquering module will select and conquer the district plan with dynamic programming method. Then, decisions will be made at the Module Three or FMCDM module while final result will be enhanced in Module Four or COD module.

Besides, there are four main data stores, which store all the information needed for the working prototype. These data stores are determined based on the selected case study and they are used to store the contour layer, license layer, river layer, and the shape optimal rulesets. All of the necessary data, which includes application dependent data and application independent data, are determined based on the forestland blocking. Each of the details is further discussed in following subsections. The detail implementation codes for all the processes in AML can be found in Appendix A.5.

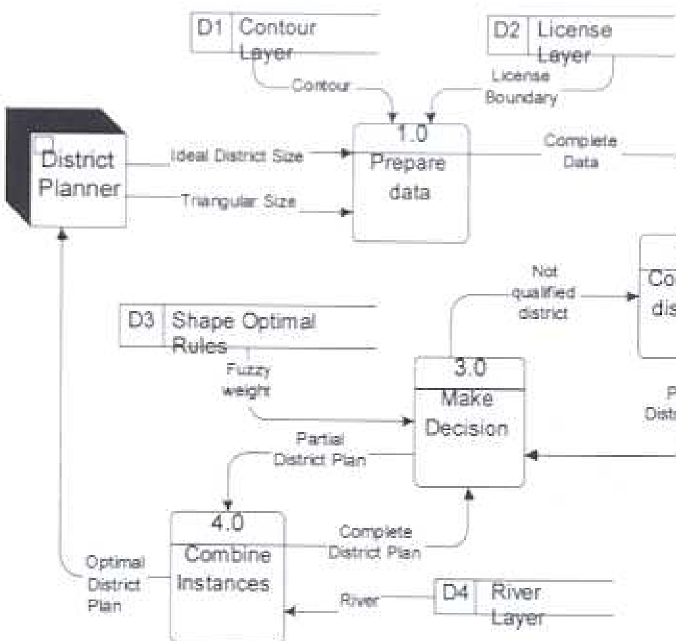


Figure 4.9: Diagram 0 of the Data Flow Diagram

4.3.1 Initialization Module

This first module is the initialization module and it has all necessary initialization processes such as definitions of all variables and their initial values for the prototyping. The entire data source and important user-defined parameters will be determined and assigned to the appropriate variables. After that, this module will start the preprocessing on all the input data so the data will be prepared into a suitable format for the redistricting process in the District Conquering module later.

The main input data needed are the data source for contour, license boundary, and river. These input data are the spatial and non-spatial data needed for the redistricting process. Besides, the district ideal size and triangles sizes are required too for the data preparation module and District Conquering module. Weights values for EM (Euclidean Measure by area-perimeter ratio), FD (Fractal Dimension), application dependent criteria (district size), degree of confidence and attitude to risk are needed too for the FMCDM module. All the inputs are then stored in the global variables for the ease of manipulation.

The necessary spatial data for the research case study are digitized topographic map with 1:50,000 scale. Contour with 100 feet interval is the basic requirement to provide the slope data. Slope in degree acts as the application dependent data requirement or criteria for the redistricting process in the prototype development. As the unit used in the prototype is in meter, the contour interval would be converted to meter from feet. Then, license boundary and river are important too to act as the natural feature in the case study. Subsequently, there are three different layers used in the prototype for overlay process (see Figure 4.10). All these layers are in line features and they will be processed and prepared into appropriate data format in this module.

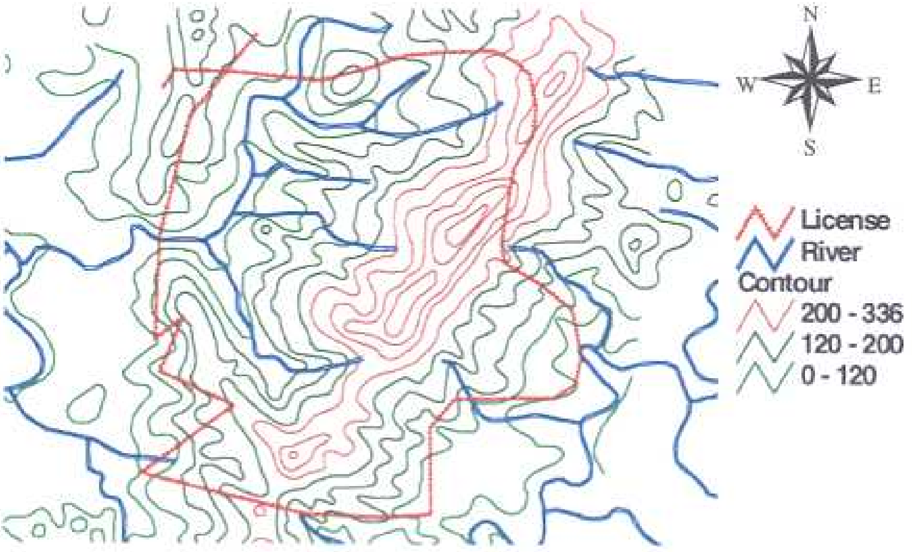


Figure 4.10: Three layers for the river, license and contour (divided into different colors according to the elevation interval) for the data source

After the prototype obtains the input data from the prototype, the module will convert the contour line into polygon coverage so it is accomplished with slope attribute that will be ready for District Conquering module. Thus, elevation point acts as input source for the case study to generate Triangulated Irregular Network data model, called TIN so that it can be then converted to polygon coverage (data format in Arc\Info). In the implementation stage, different triangle size can be used in order to test the effects of the triangle size to the district plan. Indeed, there is a textbox to get the triangular size in square meter within the input form to get the different triangles size values. The default triangle size is 500 meter square.

During this preparation process, TIN surface analysis enables the prototype to calculate slope and aspect for each triangle. The calculated values are recorded in SLOPE and ASPECT items added to the necessary data files. The slope of a surface refers to the maximum rate of change in z values across a region of the surface. The two most common methods for expressing slope are as an angular measurement in degrees, or as a percentage. For example, a rise of 2 meters over a distance of 100 meters can be expressed as a slope of 1.15 °, or as 2 percent. The aspect of a surface is the compass direction maximum rate of change in z in the downward direction. Aspect is expressed in positive degrees from 0° to 360°, measured clockwise from the north. However, only slope factor is selected as redistricting factor in the prototype development.

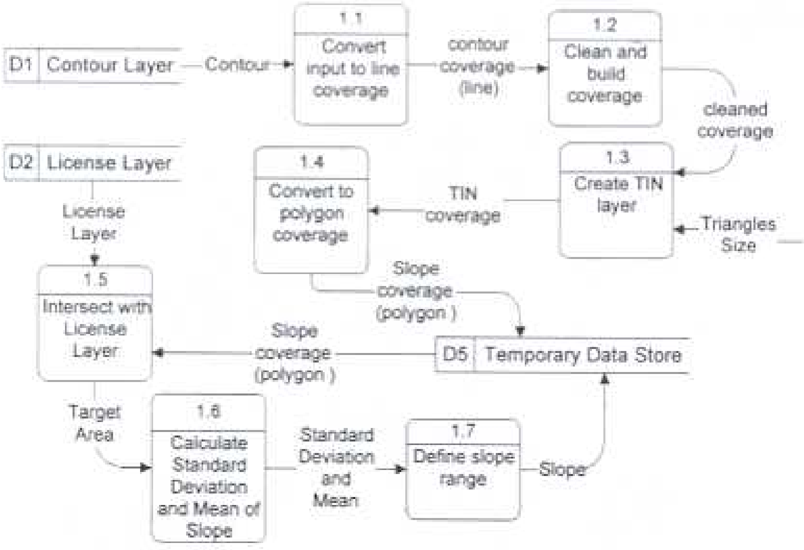


Figure 4.11: Data Flow Diagram for Module One or Data Preparation Module

The logical module structure can be illustrated in data flow diagram as Figure 4.11. There are seven important processes in order to produce and prepare spatial and non-spatial data for the prototype. The primary process is the data conversion of input source in shapefile format to contour coverage in line feature. This process ensures the input sources are in the appropriate working data format. Special command used to convert the input source in shapefile to working coverage in Arc\Info feature is shown in Figure 4.12 where the *<in_shape_file>* is the name of the source and *<out_cover>* is the new coverage name.

SHAPEARC <in_shape_file> <out_cover>

Figure 4.12: Command to convert shapefile to a working coverage in ARC\INFO

After the conversion, the line coverage is needed to generate topology to ensure the coverage is with correct arc-node (line-point) topology. Therefore, *CLEAN* command edits and corrects geometric coordinate errors, assembles node or point into lines and creates feature attribute information for each arc. Besides, *BUILD* command creates or updates a feature attribute table for the coverage.

Then, the following implementation stage is to get slope values. A transition stage of TIN coverage is compulsory to help to generate the slope value. Therefore, by using command below where the *<out_tin>* is the TIN name and *{weed_tolerance}* is the triangle size creates the TIN triangles.

CREATETIN <out_tin> {weed_tolerance}

Figure 4.13: Command to create TIN coverage based on user input

The *COVER* subcommand is used to input the source coverage, which is containing the contour lines. Then, the TIN feature is converted back to polygon coverage and slope in degree value will be calculated by command *TINARC*. The new polygon coverage with slope in degree is stored in a temporary data store for the ease of manipulation in the next module. Indeed, the mentioned slope polygon is shown in Figure 4.14.

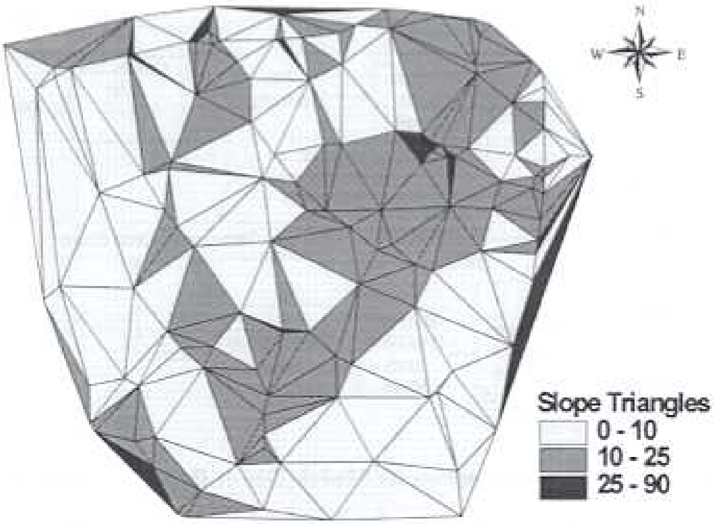


Figure 4.14: TIN triangles divided into three different slope range. Less triangles were found for highest slope range (steepet area).

During the preparation period too, the polygon coverage is intersected with the license boundary in order to get the target area. Some of the items or attributes in the polygon coverage are dropped to avoid duplication and distraction for the unnecessary data in the attribute file later. The process is proceeded with calculation of the standard deviation and the mean for the slope values to prepare interval values for each stages of DP module later and to define the slope range in this module. After the seven logical processes, the necessary spatial and non-spatial data are ready and will be passed to the following module.

4.3.2 District Conquering Module with Dynamic Programming

District Conquering module is the module that start to redraw the district boundary line for the forestland blocking based on the predefined criteria such as the slope and district area size. The smallest sub-problems called instances are prepared in the previous module or Initialization Module for the shape-based redistricting problem. In addition, the triangles are divided based on the predefined slope range into three different groups. Thus, this module will use the smallest sub-problems that are the smallest triangles with the slope features and proceeds with Dynamic Programming method to solve the problem from the last instances to the earliest one. All of the unsatisfactory result from one instance will be passed to another until it reaches optimality.

There are seven important processes to create a partial district plan as shown in data flow diagram in Figure 4.15. This module is called three times from the highest slope range to the lowest one in order to get three partial district plans, which will then be combined to the final district plan in the Combine Optimal District (COD) module. The definition of the necessary data is ready from the Initiation module and the combination process is conducted in the COD module. The processing scheme in this module is based on the dynamic programming method. Subsequently, the detail process in this module is mainly conducted to create the district plan based on the SELECT ADJACENT method.

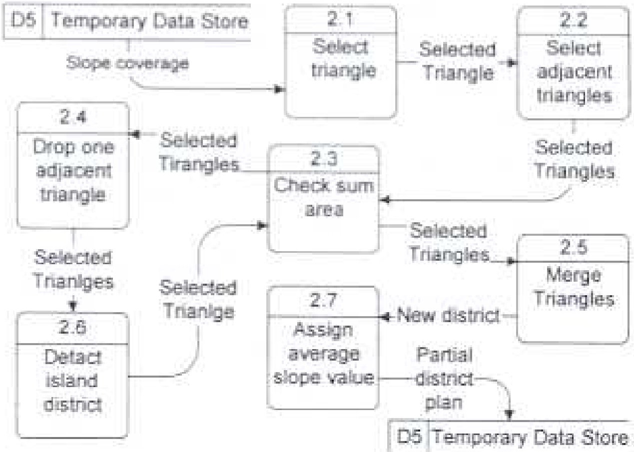


Figure 4.15: Data Flow Diagram for Module Two or DP module

The preliminary process in this module is to select or obtain the polygon coverage with specific slope values from the temporary data store, which stores the ready-to-use information from Initialization module. ARCEDIT is ARC/INFO's program for editing coverage coordinate and attribute data. This program is used for selecting polygon features by specifying the coverage name with the *EDIT* command and the feature class with the *EDITFEATURE* command. The algorithm for implementing this module can be found in Figure 4.16 and each of the steps is developed based on the design in the data flow diagram in Figure 4.15. Then, the algorithm is converted to AML script in ARC/INFO in a module as shown in Appendix A.5. In addition, there are some clearer justifications in the following part to explain the important steps and command used in the implementation process.

```
1.  Select all polygon
2.  Calculate number of polygon selected, n
3.  Do on each polygon, x until n
    3.1  select each polygon, x
    3.2  select adjacent polygon to the x
    3.3  count the number of polygon selected, y
    3.4  if x <> y
         3.4.1     if Sum of Area(selected polygon) < Area(ideal district size)
              3.4.1.1   if Sum of Area(selected polygon > Area(ideal district size)
                        3.4.1.1.1 drop the polygon with largest area
                        3.4.1.1.2 Go to 3.4.1
         3.4.2     else merge all polygon
4.  Loop
```

Figure 4.16: Algorithm to implement the Districting Conquering module

In the first process, the polygon is selected based on the internal record number *($Recno)* by *SELECT* command. Logical operators like *AND* and *OR* are used in order to avoid unnecessary selection. Besides that, during the selection process, *RESELECT* helps to select a subset of the currently selected set of features and *UNSELECT* removes selected features from the current selected set. On the other hands, *ASELECT* adds more features to the currently selected set. The following command is extremely important in order to select the adjacent polygons.

```
aselect %cover% POLY ADJACENT 0
```

Figure 4.17: Command to select adjacent polygons

However, the command above can only be used within ARCPLOT. Therefore, it was necessary to transfer the ARCEDIT selected set to the ARCPLOT environment with the *SELECTPUT* command so that selecting the adjacent polygon can continue the process. Thereafter, the ARCPLOT selected set is transferred back to the current edit coverage in ARCEDIT with the *SELECTGET* command. Then, *SHOW* command is used to see the number of polygon features selected while *DRAWSELECT* helps to draw all currently selected features with the symbol specified in the *SETDRAWSYMBOL* command and controls whether selected features will be highlighted as they are selected.

In the program, *&IF*, *&THEN*, and *&ELSE* allow the statements to be executed conditionally. Besides, iterative process is conducted by *&FOR* loops. *&DO* ... *&END* blocks allow multiple lines to be conditionally executed based upon a single test. Besides, this independent operation is created through routines. The routine alters the order of execution by transferring control to a different statement in the program. *&CALL* transfers control to the statement following *&ROUTINE* and *&ROUTINE* delimits the beginning of a routine block. The end of the routine block is either *&RETURN*, *&STOP*, or the end of the file.

For each slope range, each polygon is selected to find the adjacent polygons. The conquering process is to merge the polygons if the total area of the selected polygons is not exceeding the allowed area. If it is exceeding the area range, one polygon is dropped each time until it gets the total area reach to the allowed district area. The allowed or ideal area is based on the global variable that was included in the input form at the beginning stage.

MERGE SELECT is the command to combine selected features into a single feature. The merged polygon would not retain any attributes associated with any previously selected features unless the program calculated and assigned values to an item by the command below:

CALCULATE <target_item> = <arithmetic_expression>

Figure 4.18: Command to assign value to an item

The process ends by saving the *SAVE* command so that it will clear the selected set and clear the transaction list. Appendix A.3 has the detail implementation codes of the merge process. For the dropping process, total area for adjacent polygons needed to be summed up. Therefore, *STATISTICS* command helps to calculate summary statistics for the values of the currently selected features set as the following subcommands below.

SUM <item>
MEAN <item>
MINIMUM <item> or MIN <item>
MAXIMUM <item> or MAX <item>

Figure 4.19: Subcommand in the *STATISTIC* command

Then, another command would be used to display the list of unique item values in order to store the temporary value as below:

[LISTUNIQUE <specifier> -INFO <item_name> {output_name}]

Figure 4.20: Command used to display the list of unique item value

Besides, the absolute value of the calculated area and the result of the calculation of an ARC expression will be returned by the AML language as below:

[ABS <number>]
[CALC <arc_expression>]

Figure 4.21: Another two commands used to get the caculated value

The recursive function is conducted by calling the function itself in the routine. Therefore, the routine stops by passing the control to the statement following *&LABEL* having the same label. It is controlled by the *&GOTO <label>* command. The detail implementation codes for the whole module is shown in Appendix.

4.3.3 Fuzzy Multiple Criteria Decision Making (FMCDM) module

The objective of this module is to obtain the Integrated Compactness Indexing (ICI) based on the developed algorithm in Section 3.3.3. These ICI are extremely importance to act as the decision indexing to determine the optimality attained by the enhanced algorithm. The compactness indexing produced from each of the selected compactness measurement criteria and the criteria of area size will incorporate together with the weighing vector contributed by the DM on each criterion. The ICI value will consider the control variable on the confidence level and attitude to risk from the DM. In short, the main task for this module is to produce a meaningful and useful indexing called ICI, which represents and analyses the shape-based redistricting process in a natural as well as systematic way.

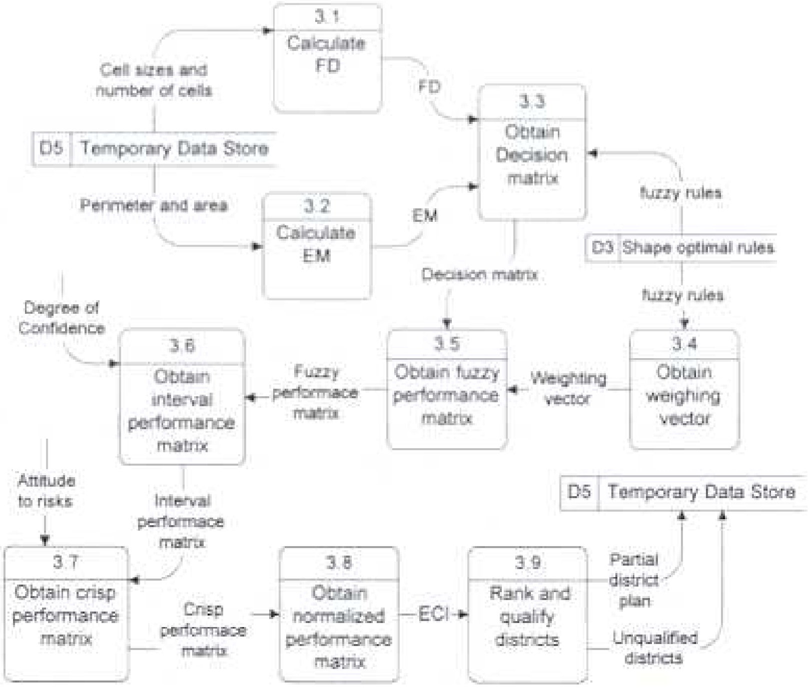


Figure 4.22: Data Flow Diagram for Module Three or FMCDM module

The logical scheme for the implementation is shown in Figure 4.22. There are nine processes involved in order to produce the Integrated Compactness Indexing (ICI) for each district. Firstly, the partial district plan with appropriate parameters is passed to the initiate processes (Process 3.1 and 3.2) to produce each compactness indexing from Fractal Dimension and Euclidean Measure based on area-perimeter ratio based the discussion in Section 3.3.2.

Then, the decision matrix is defined according to the shape optimal rules that defined in Section 4.2.2. At the same time for the definition of the decision matrix, the enhanced model needs to obtain the relative weight value for each of the criteria based on the fuzzy rules from shape optimal rulesets. It is to indicate its importance for compactness rule based on the efficiency, stability, and consistency of the methods as mentioned in Section 4.2.1. Later, both the decision matrix and weighting vector is aggregated into performance matrix. The fuzzy performance matrix will be aggregated with the degree of confidence entered during the initiation time in order to get the list of interval performance matrix for each of the district. The newly calculated matrix is next aggregated with the attitude of risk that is one of the global variables assigned at the beginning stage to get crisp performance matrix. After that, normalization is done in order to get the normalized performance matrix with the ICI values for each of the district. Finally, the district is ranked and qualified according to the ICI into qualified partial district plan and unqualified district. Both of the output will be passed to Combine Optimal District module to integrate all the partial district plans from each stage in Conquering District module.

4.3.4 Combine Optimal District Module

The main tasks in this module are to integrate all the partial district plans and to take into account the river layer. In other words, the partial district from the previous module will be combined back to the original district plan before it is intersected with the river layer. Therefore, there are three main processes for these tasks as shown in data flow diagram in Figure 4.23.

At first, the data is passed from the temporary data store to the first process to combine the partial district plan with the command as in Figure 4.24. It is to get coverage result from the same slope range (*<upate_cover>*) and update the result with the original coverage (*<in_cover>*). This command replaces the input coverage areas with the update coverage polygons using a cut-and-paste operation.

After the integration, the complete and original district plan is overlaid with the natural boundary. Consequently, *INTERSECT* command in Figure 4.25 is used to intersect the river coverage with complete district plan to get river within the necessary area. The details for the implementation of the algorithm in the prototype are described in the Appendix.

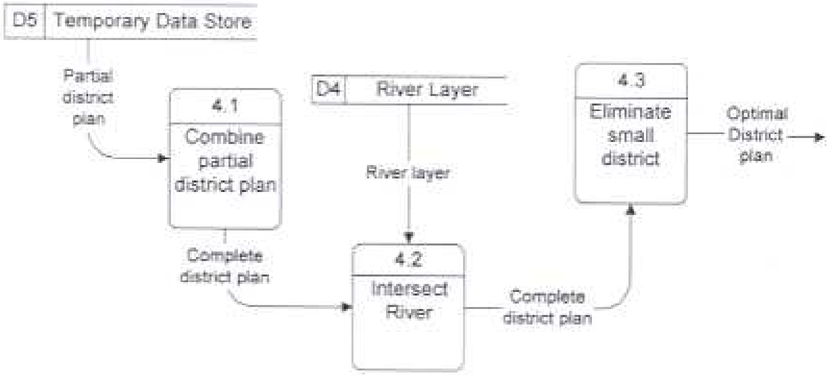


Figure 4.23: Data Flow Diagram for Module Four or Combine Optimal District Module

UPDATE <in_cover> <update_cover> <out_cover> POLY KEEPBORDER

Figure 4.24: Command used to combine and update the coverage

INTERSECT <in_cover> <intersect_cover> <out_cover> LINE

Figure 4.25: Command to use *INTERSECT*

4.4 Summary

The implementation of the enhanced model for the shape-based redistricting based on the Integrated Compactness Indexing using Fuzzy Multiple Criteria Decision Making through Dynamic Programming was done through a prototype in this section. The implementation of the prototyping is a demonstration of the practicality for the conceptual design of shape-based redistricting algorithm. Before the implementation, the primary goal and the affected factors must be considered carefully for incorporating in the Fuzzy-AHP methods. Subsequently, the selected goals and criteria will help effectively to define the optimal rules. The optimal rule in this research is heavily concerned on the shape compactness because of these research objectives. Therefore, two compactness measurements are selected as two involved criteria. The effectiveness of these criteria must evaluated and surveyed through a preliminary experiment with numerous shape conditions because they act as a proof to define the fuzzy rules on the district shape optimality. However, the determination of these district shape optimality rules is not simple because of the subjective and vague result. Consequently, a testing on this definition of the district compactness rule is important to measure their graduate changes and inconsistency in the following chapter. Besides, the weighing vectors contributed by the DM plays an important role to weight the importance of the selected criteria. Thus, the DM need to certain on the usefulness and the contribution of the criteria to the redistricting process. Anyway, the use of the variables about confidence level and attitude to risk will particularly assist the DM during their decisions making process on the weighing vector. In other words, the Integrated Compactness Indexing (ICI) produced from the Fuzzy-AHP is incorporated with the nature of the real decision making process. Lastly, the supportive tool in GIS facilitates the maintenance and manipulation of the spatial and attribute data. Though the prototype seems to be simple and easy to be implemented, the design and implementation cause many efforts.

CHAPTER 5: RESULT AND DISCUSSION

This chapter discusses on the result and analysis of the shape-based redistricting prototype and the uses of an Integrated Compactness Indexing with Fuzzy Multiple Criteria Decision Making based on multiple compactness measurement into redistricting technique. Analysis and evaluation for the developed shape-based redistricting algorithm aims to analyze the relationships of all the variables involved and the effects of different type of data toward the prototype. There is performance comparison between the district plan with and without the enhanced model. These analysis and comparisons are important to assess the degree to which the prototype conforms to the design and specification. Besides, it helps to assess the degree of optimality of the district plan generated from the enhanced model. In brief, the result and analysis focuses on the applicability, effectiveness, and limitations towards the Integrated Compactness Indexing (ICI) and the enhanced shape-based redistricting algorithm.

The overall performance of the developed shape-based redistricting algorithm is analyzed using statistical methods that include Mean, Mean Deviation, Maximum, Minimum and their Difference of compactness indices for each district to evaluate compactness of a district plan. Indeed, a plan's compactness is defined as the mean compactness of its districts used in Iowa and in Michigan in United State for political redistricting (Altman, 1998). Besides, other research mentioned that its least compact district determines the compactness of a plan. These statistical methods will be applied at different conditions for the developed shape-based redistricting algorithm in order to find out its performance under different circumstances.

There are three main areas of analysis and evaluations carried out on the redistricting algorithm:

(a) result and analysis with respect the different input variable like interval to calculate slope and the restricted boundaries such as river and license boundary.
(b) result and analysis on the advance FMCDM component,
(c) performance evaluation and comparison on district plan produced with and without the enhanced redistricting algorithm,

In order to delimit unnecessary confusion, the phrase 'Enhanced Redistricting Algorithm' mentioned in the following section means the specific shape-based redistricting algorithm developed in this research.

5.1 Result and Analysis for the Shape-based Redistricting Prototype

An input form is used to get necessary input data for this prototype as in Figure 5.1. The input form consists of textbox and optional selection for user especially to input available data sources. Lists of optional values in the interface are for the selection of the desired values and defaults values. Three listbox are provided to allow the users to select the three different layers such as contour, license boundary, and river as requirement for prototype. Besides, users have freedom to enter their desired and specific name for the output layer. Another two important variables needed for the Enhanced Redistricting Algorithm are the triangle size and the ideal area district size in meter unit. All the necessary variables will be passed through each module to produce an optimal district plan.

This section aims to analyze and evaluate the performance of the Enhanced Redistricting Algorithm in respect with different type of the input data. First, the evaluation is on different triangle size and their relationship with the ideal district size to generate the district plan. Then, the evaluation is on different shapes of license boundary on affecting the compactness of the district plan generated from the prototype. With the evaluation conducted in this section, this research will confidently handle the Enhanced Redistricting Algorithm.

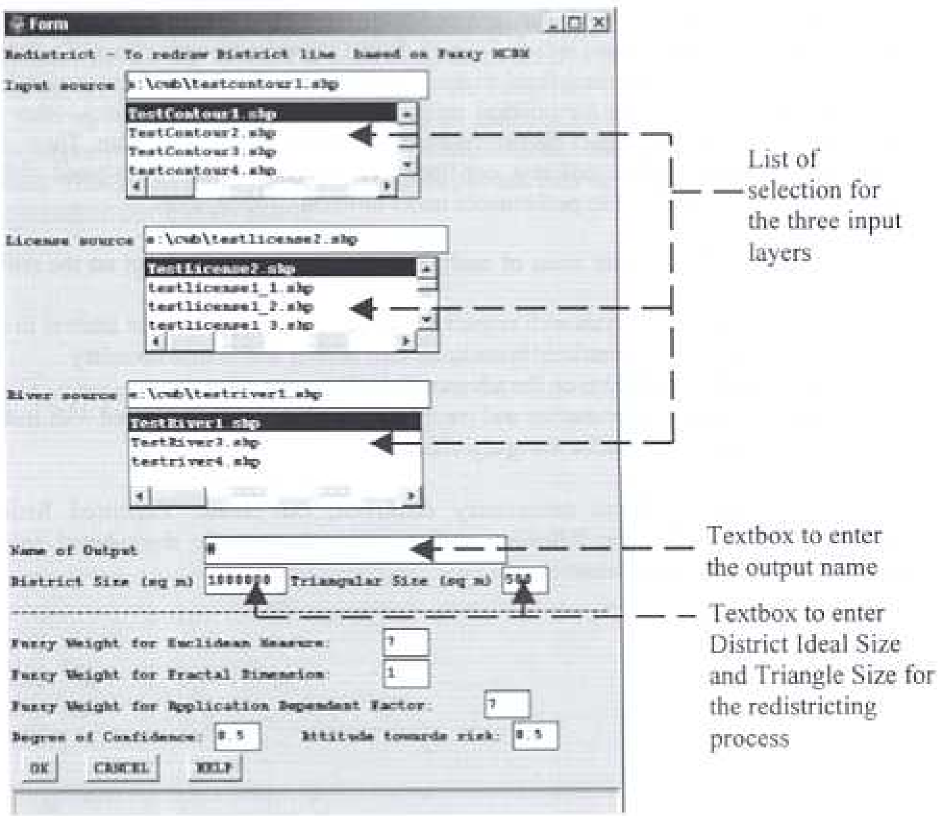
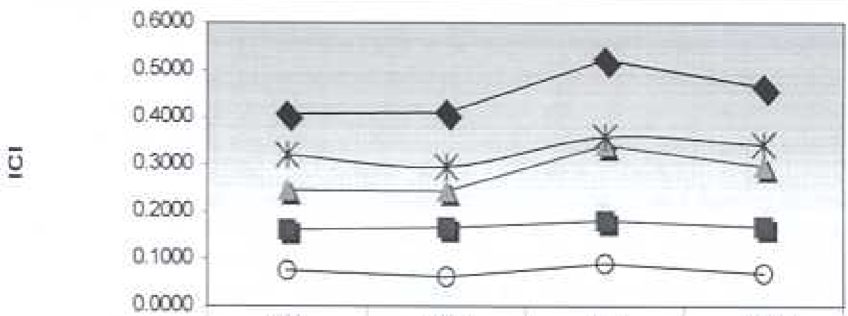


Figure 5.1: Program Interface for Input form

5.1.1 Result and Analysis with Respect to Different Triangles Size

The smaller the triangle size, the longer is the processing time for the Enhanced Redistricting Algorithm in the prototype to generate an optimal compact district plan. However, different triangle size will produce different district plan output because the larger the triangle, the generated districts in the district plan will have a rough and not detail boundary line. A suitable triangle size is important to enable the prototype to perform effectively and in a promising processing time to generate a district plan. Therefore, a series of input is tested for different triangle sizes as the input variables in order to analyze their effects to the algorithm. A survey of four different sizes of the triangles with 500, 1000, 2000 and 5000 meter square are used to produce a 1,000,000-meter square district plan. This research surveys the output or the district plan and draws charts for the statistical measurement that includes Maximum, Minimum, Difference between them, Mean and Mean Deviation of the Integrated Compactness Indexing (ICI) for the district plan as shown in Figure 5.2.

According to the result produced by the Enhanced Redistricting Algorithm, the output that is the district plan is influenced by the triangle size. The triangle size of 1000 that is one over one thousand ratio from the total district plan produces the best relative result because the Mean Deviation among the districts is only 0.0621 compared to the rest of the result based on other triangle sizes. However, the Mean value for this district plan is 0.2949 only, which is the lowest compared to the other district plans. Thus, triangle size of 5000 is considered as better result because its result is with the Mean Deviation of 0.069 and Mean value of 0.3425. Although the Mean value is higher than the district plan based on triangle size of 1000, the Mean value of this plan is much higher than that. Therefore, the considerable district plan is based on triangle size of 5000 or 0.5 percent from the district plan. Another important aspect that influenced by the triangle size value is the working time of the enhanced algorithm to produce the district plan. The smaller the triangle size, the longer working time is necessary to produce the district plan. Therefore, triangle size of 5000-meter square or 0.005 of the size of district plan is the optimum value to produce a 1,000,000-meter square of the district plan.



	500	1000	2000	5000
Maximum	0.4071	0.4105	0.5228	0.4646
Minimum	0.1614	0.1664	0.1813	0.1681
Difference	0.2457	0.2441	0.3415	0.2965
Mean	0.3197	0.2949	0.3591	0.3425
Mean Deviation	0.0754	0.0621	0.0902	0.0690

Triangles Size

Figure 5.2: ICI for Different Tirangles Size

5.1.2 Evaluation with Respect to Restricted Boundary

Redistricted boundaries such as political boundary like license boundaries and river will also affect the result of the Enhanced Redistricting Algorithm especially in compactness aspect because the restricted boundary always causes non-continuity. Therefore, some compactness standards like in IOWA General assembly in 1980, Bill generally defined compact district shape to the extent permitted by the restricted boundary (Altman, 1998).

Consequently, this research tests and analyzes the influence of the restricted boundary to the Enhanced Redistricting Algorithm with five different license boundaries with different shape such as rectangles, circles and some other odd shape as shown in Figure 5.3. Later, a chart is produced to represent to redistricting output as in Figure 5.4. The result shows that rectangle (V3) and circles (V4) gave better result compared to the other three license boundaries because they give highest Mean value and lowest Mean Deviation value. The Difference between the Maximum and Minimum ICI value for these two type of license boundaries which give the lowest result proves the fact that rectangle and circles shape of license boundary will produce the optimum compact district plan. In short, shape of the license boundary will affect the district plan due to the normalization in enhanced algorithm. As relative measure of individual district is important in assessing the compactness of the district plan, the license boundary must be considered as one of the evaluation aspects.

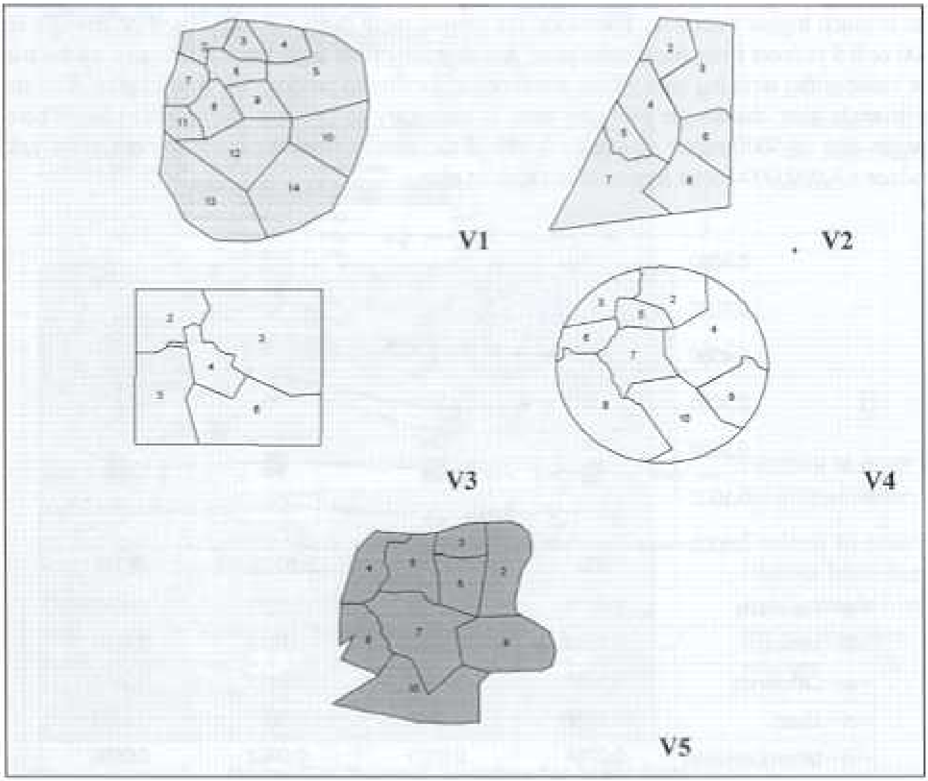
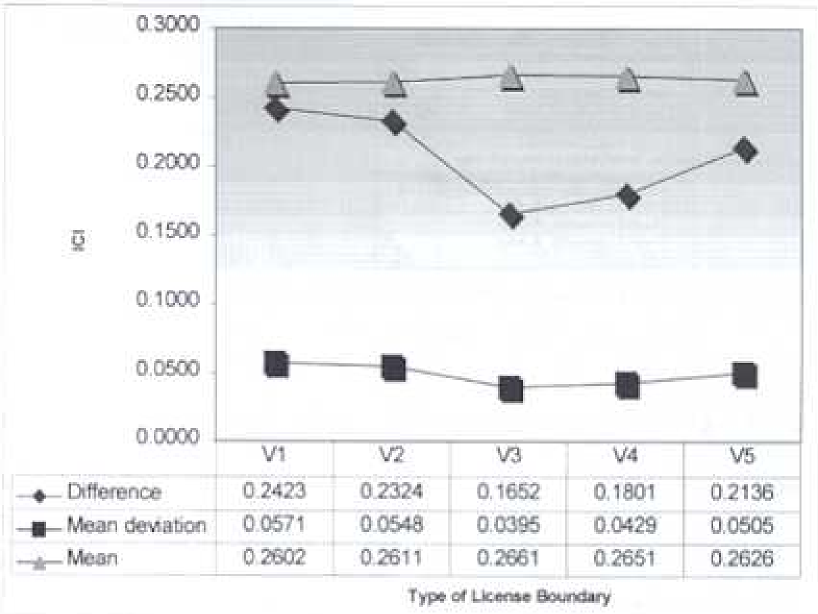


Figure 5.3: Difference shape for the selected License Boundary



	V1	V2	V3	V4	V5
Difference	0.2423	0.2324	0.1652	0.1801	0.2136
Mean deviation	0.0571	0.0548	0.0395	0.0429	0.0505
Mean	0.2602	0.2611	0.2661	0.2651	0.2626

Figure 5.4: Result for different License Boundary

5.2 Performance Evaluation of Fuzzy Multiple Criteria Decision Making

FMCDM is the underlying concept that supports the Integrated Compactness Indexing (ICI) in the Enhanced Redistricting Algorithm. Thus, it plays an important role in the enhanced algorithm because performance for the usage of this method in the ICI will give an impact to the algorithm. Indeed, this research will analyze the performance of the ICI for the variety of the Linguistics Terms' definition of the decision matrix and weight vectors, the observation parameter for attitude to risk, α and level of confidence, λ of the district planner. Besides, the evaluations are conducted regarding number of considered criteria and the decision indexing or the ICI at each stage of decision making process. Most of these variables are user-defined in the textbox in the program interface as shown in Figure 5.5 and selection of values can be found in a listbox once users right-click on the textbox. The details on the evaluation conducted in the FMCDM are discussed in the following sections and the triplets commonly used in the following section shows the importance or weight of each criterion among three criteria in the prototype. Criteria one refers to the Area Size, criteria two refers to the Fractal Dimension and criteria three refers to the Euclidean Measures base on area-perimeter ratio. Besides, the district planner is called as the DM in the following sections.

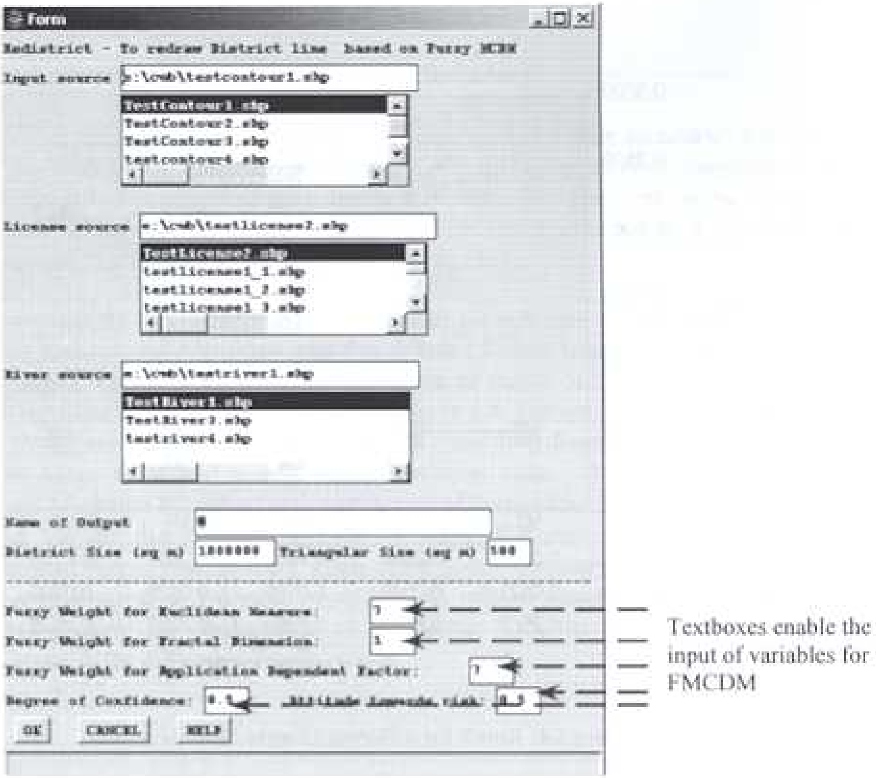


Figure 5.5: Program Interface for Input form

5.2.1 Linguistic Term Definition in the Shape Optimal Rules

The ICI in prototype is developed by the definition of linguistics terms based on circle but the assessment of compactness in district plan is with the relative comparisons among each district in the entire individual district plan. Therefore, the most compact district can be categorized as octagon but not the circle shape. Therefore, the research studies the compactness indexing with different types of definitions for the Shape Optimal Rules based on circles, octagon, hexagon, and rectangle to compare and contrast their effects to the enhanced algorithm. The result in Table 5.2 shows that circle still is the most compact one when they are used alone due to their smallest value in the comparison for both the Euclidean Measure based on area-perimeter ratio (EM) and Fractal Dimension (FD). However, the value of compactness indexing for octagon and hexagon are extremely close with circle and this shows that the effects of their definition of the linguistics term according to circles, octagon, or hexagon to the overall Enhanced Redistricting Algorithm will not very obvious. However, the effect of rectangle is apparent due to the higher compactness indexing for EM.

Table 5.2: Compactness for different type of polygon by using EM and FD.

	EM	FD
Circles	12.5779	0.9994
Octagon	13.2907	1.0102
Hexagon	13.8817	1.0107
Rectangle	16.0017	1.0074

In order to show the seriousness of the linguistic terms definition at the Shape Optimal Rules toward the overall performance of the Enhanced Redistricting Algorithm, three different sets of the linguistic terms' definitions for the Euclidean Measure criteria is developed. Figure 5.6 shows the three sets of Linguistic Terms' definition. These sets of definition is combined with the same Linguistic Terms for the other criteria in order to see the influences of a particular criterion on the enhanced algorithm. Therefore, the research evaluates the same input data by using the newly defined linguistics terms for the combination of weight value by 717. It means that the research assigns the first criteria with weight $\bar{7}$, second criteria with $\bar{1}$ and third with $\bar{7}$.

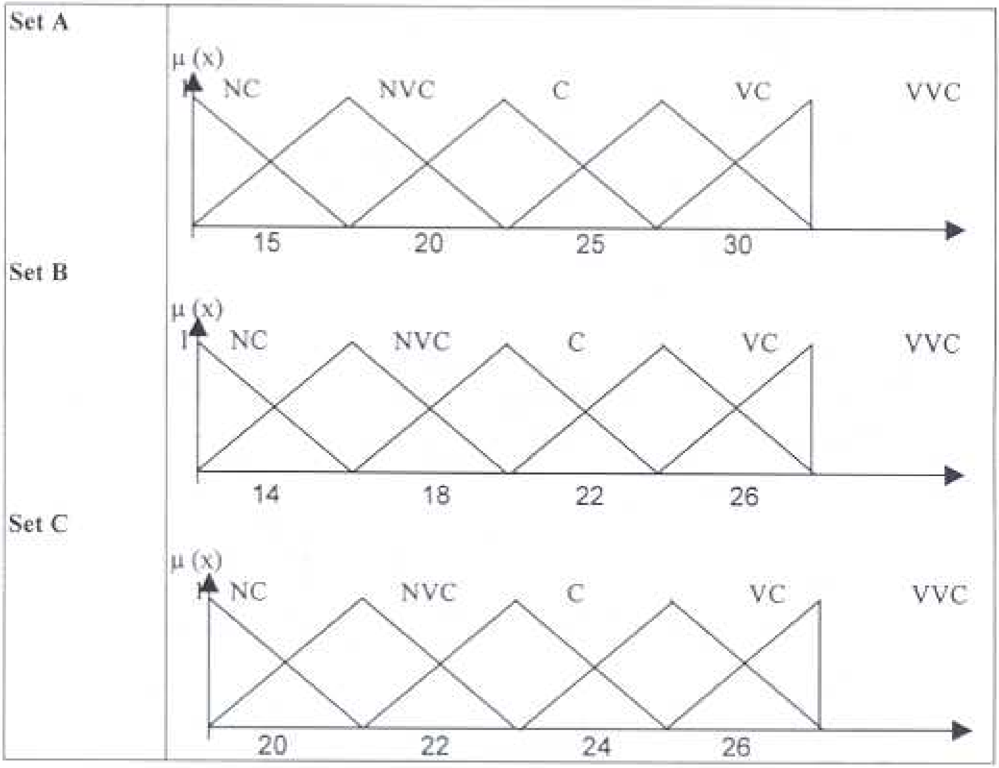
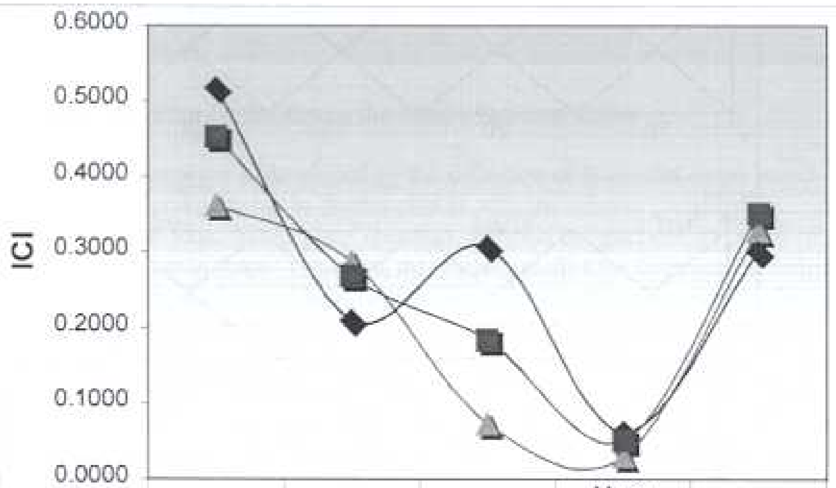


Figure 5.6: The three tested different sets of Linguistics Terms' definition for EM

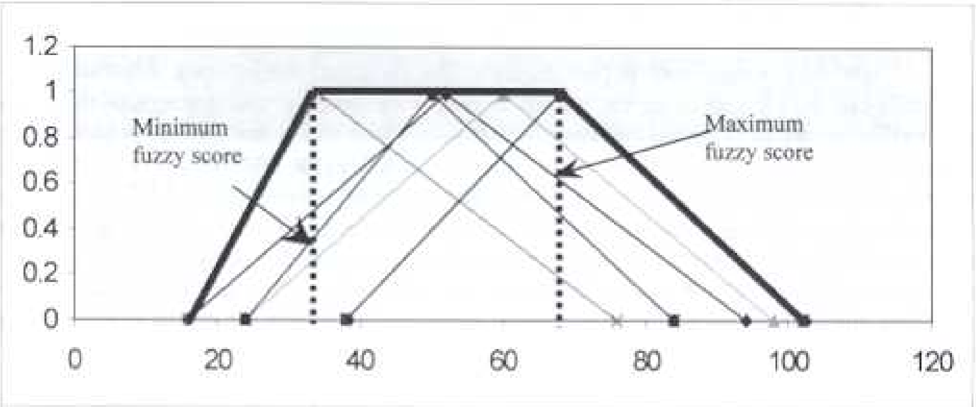
From the Linguistics definition, the research is able to produce the membership function of the triangular fuzzy number for all three sets of definition in Figure 5.8. Color lines in the Figure 5.7 shows that the membership functions for each district whereby the bold line is the overall membership function for the district plan. The larger the values of the fuzzy scores, the more compact of the district plan with respect to the criteria items. Therefore, the definition in Set C produces better result because the membership function is the largest. Most of the districts in Set C gives higher fuzzy scores (up to 90) compare to the other two sets. Although Set B gives high fuzzy scores (up to 100), most of the district gives the smallest fuzzy scores which is less than 20. Therefore, the research can conclude that the Deviation value in Set C will be larger. On other hand, Set B gives average result with moderate fuzzy scores with respect to the criteria item.

In order to have clearer view, the research then calculates and obtains a set of statistical value to compare the result of the three different sets of linguistics term definition in Figure 5.7. The Maximum value for Set B is the highest and the Minimum value is the lowest compare to the Maximum and Minimum values for other sets. Therefore, the Difference is large and the Deviation is big, Mean value very low. For Set A, Maximum value is the lowest and the Minimum value is the highest. Although the Difference between Maximum and Minimum value and the Mean Deviation value is lowest, its Mean value is average. On the other hand, Set B has average Maximum and Minimum value. Therefore, the Difference and Mean Deviation are average as well. However, the Mean value is the highest one among all sets of definitions. Thus, Set B has give better result and the different between all three different sets of linguistics term definition give quite small differences. In conclusion, shape optimal rules based on circles, octagon and hexagon are better than rectangles.

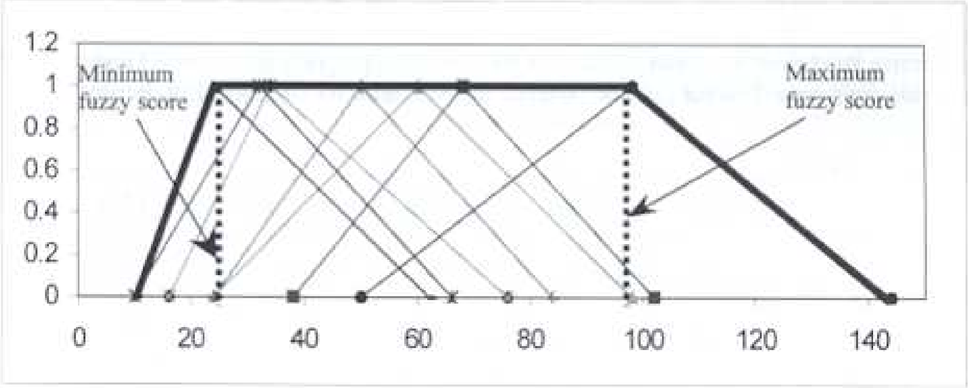


	Maximum	Minimum	Difference	Mean Deviation	Mean
Set A	0.3619	0.2876	0.0743	0.0306	0.3283
Set B	0.5173	0.2097	0.3076	0.0633	0.3008
Set C	0.4532	0.2692	0.184	0.0519	0.3506

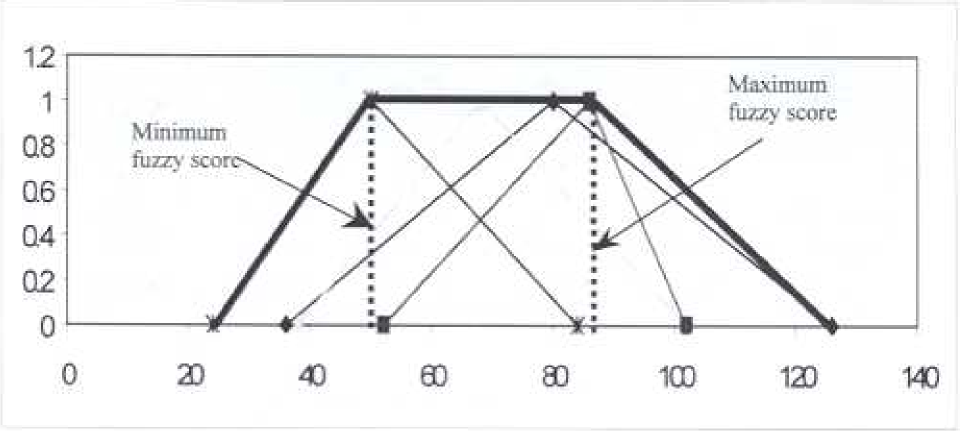
Figure 5.7: Statistic for Set A, B and C



Set A



Set B



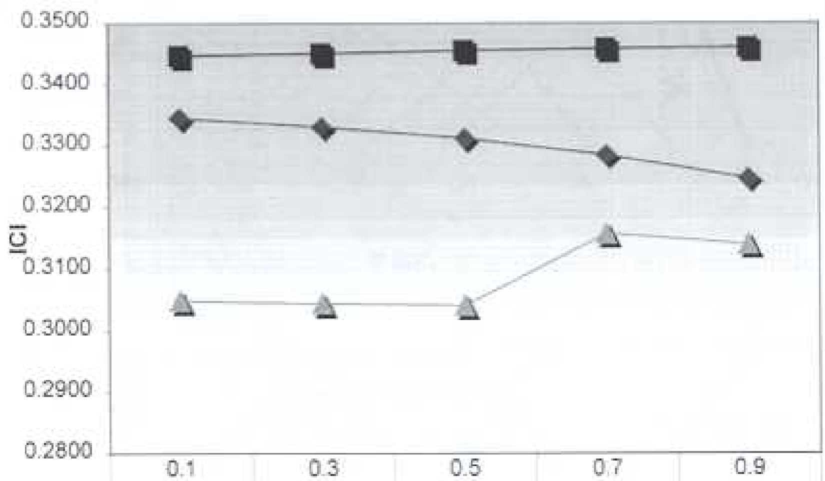
Set C

Figure 5.8: Membership function of triangular fuzzy number for District Plan based on different set of linguistic terms' definition

5.2.2 Observation Parameter, α and λ

One of the significant improvements of the Enhanced Redistricting Algorithm is the availability of the parameters on the attitude to risk, α and level of confidence, λ of the district planner. These parameters act as observation parameter, which aim to analyze the decision making behavior of the decision makers. Therefore, the research defines the DM based on their attitude to risk into three different groups that are the optimistic, moderate or pessimistic DM. For each group of the decision makers, the research let the level of confidence, α = 0.1, 0.3 , 0.5, 0.7 and 0.9, then the research can obtain the overall performace indexing or integrated compactness indexing for alternatives of district plan based on different combination of weight values. Following that, the algorithm enables the ranking of the alternative correspondingly from the results in Figure 5.9 to Figure 5.11 respectively.

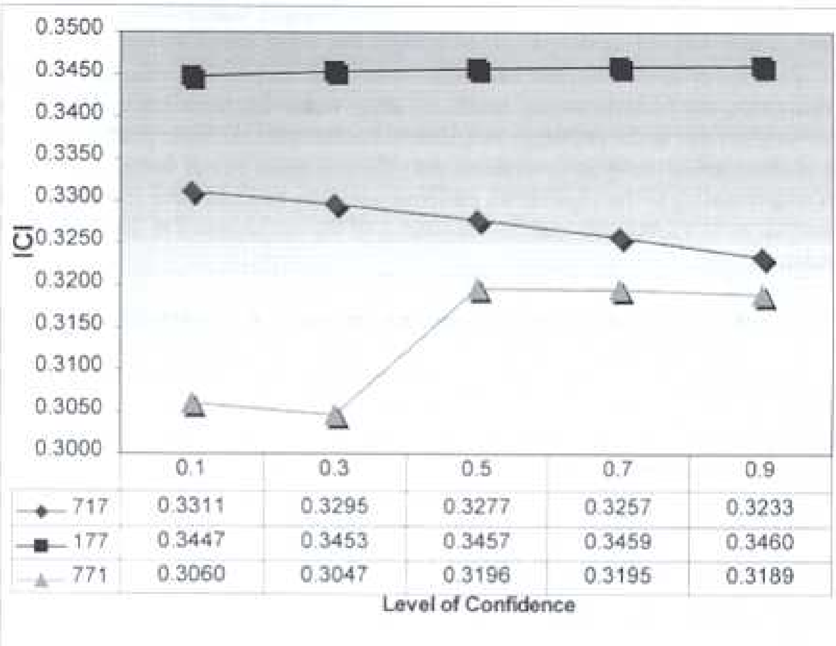
The results in Figure 5.9 to Figure 5.11 show that district plan based on weight value 177 for the three selected criteria is clearly the best choice under almost any degree of confidence of the decision makers with various attitudes towards risk. It is clear that these observation parameters can adequately reflect the uncertainty and impressions associated with the DM's subjective judgement in human thinking. It provides an appropriate tool to better understand the decision problem and his/her decision behavior to obtain effective and consistent decision.



	0.1	0.3	0.5	0.7	0.9
717	0.3345	0.3331	0.3312	0.3285	0.3246
177	0.3447	0.3451	0.3455	0.3459	0.3460
771	0.3048	0.3045	0.3041	0.3158	0.3139

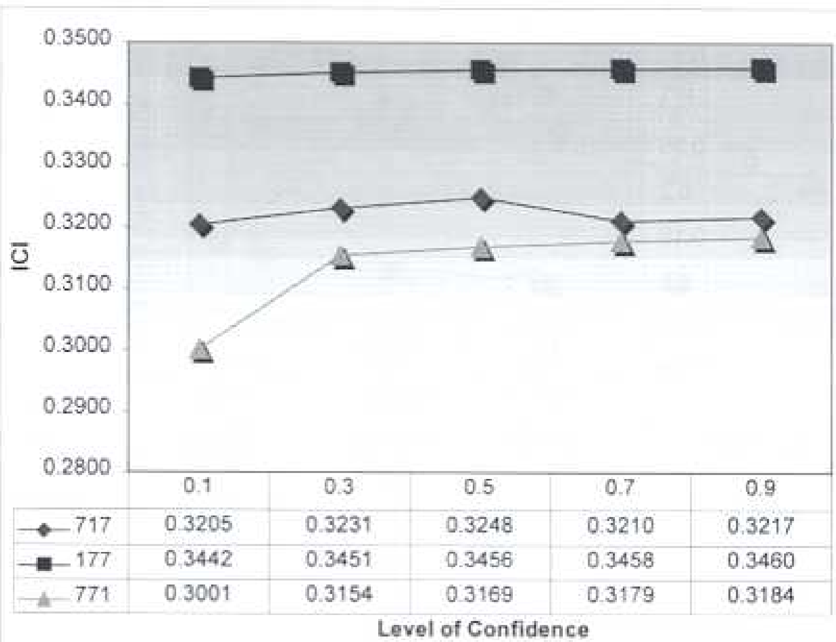
Level of Confidence

Figure 5.9: Mean of ICI for Optimistic



	0.1	0.3	0.5	0.7	0.9
717	0.3311	0.3295	0.3277	0.3257	0.3233
177	0.3447	0.3453	0.3457	0.3459	0.3460
771	0.3060	0.3047	0.3196	0.3195	0.3189

Figure 5.10: Mean ICI of Moderate



	0.1	0.3	0.5	0.7	0.9
717	0.3205	0.3231	0.3248	0.3210	0.3217
177	0.3442	0.3451	0.3456	0.3458	0.3460
771	0.3001	0.3154	0.3169	0.3179	0.3184

Figure 5.11: Mean ICI of Pesimistic

5.2.3 Number of Compactness Rules and their Weight Value

The number of the compactness measurement and their weight values play an important role to determine the optimal output result. In other word, the three criteria considered with respective weight value in the prototype help determine the optimal district plan. Thus, this section conducts another useful analysis to evaluate the effect of these values toward the prototype for thorough understanding of the algorithm's performance. One, two and three criteria are considered for the analysis in this section to analyze the change of the output result to the gradually increase weight value.

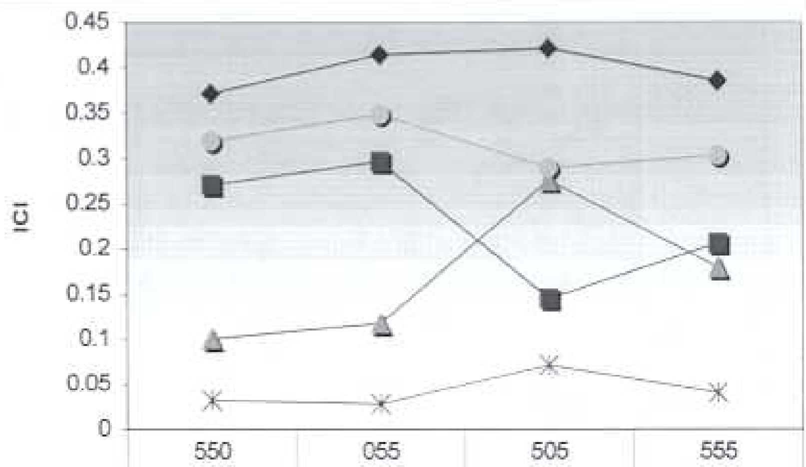
The analysis will start to consider either two criteria with equal weight values to compare with the output generated from the prototype with the consideration of three criteria. The criteria are not considered when the weight value equal 0. Therefore, the analysis consists of the following combination:

a) Euclidean Measure with Fractal Dimension, (triplet = 055)

b) Euclidean Measure with District Area Size. (triplet = 505)

c) Fractal Dimension with District Area Size. (triplet = 550)

Figure 5.12 shows the performance of the each of the weight combination towards each of the combination.



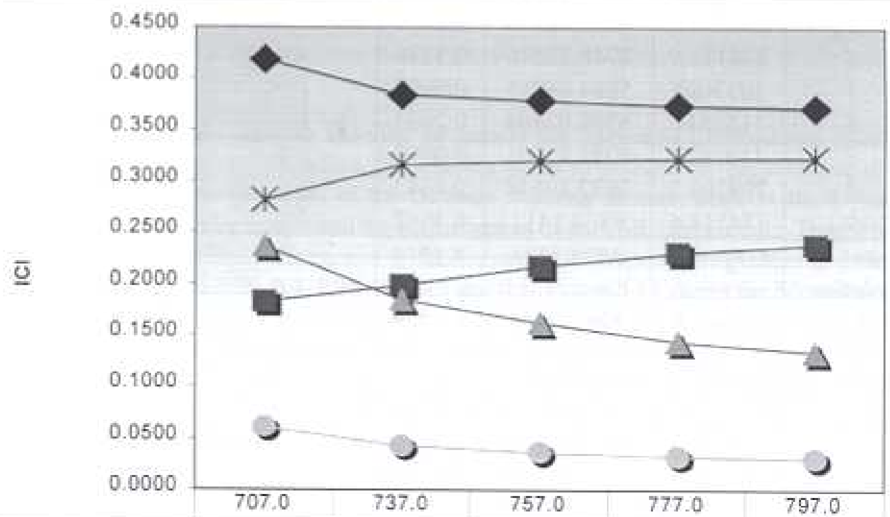
	550	055	505	555
Maximum	0.3716	0.4142	0.4212	0.3852
Minimum	0.2711	0.2967	0.1452	0.206
Difference	0.1005	0.1175	0.276	0.1792
Mean	0.3194	0.3482	0.2897	0.3034
Mean Deviation	0.0322	0.0284	0.0710	0.0405

Weight

Figure 5.12: Result for different weight value to test number of considered criteria

Without criteria two or Fractal Dimension, the prototype produces the worst output result because it has the lowest Mean value and highest Mean Deviation. In other words, without the Fractal Dimension, the district plan is not very compact and the boundary of each individual district is complex. On the other hand, without Euclidean Measure and District Area Size criteria, the prototype performs quite well because of lower Mean Deviation value and higher Mean value. However, the Enhanced Redistricting Algorithm without the District Area Size criteria produces better result between both of these conditions. When three of the criteria are assigned with the same value, the result is fine and the algorithm considers all the criteria at the same time.

It is clear that the consideration of Fractal Dimension will improve the performance of prototype by giving optimal result according to the analysis. Therefore, second analysis is conducted to analyze the effects for FD as control variable to the other criteria by giving same and specific weight values for the other two criteria but gradually increase weight value for FD in order to analyze its performances of the prototype for this aspect. Therefore, the analysis is on the same set of input data with same weight value for District Area Size and Euclidean Measure but increasing weigh value of triplet from 707, 737, 757, 777, 797). Figure 5.13 shows the result performances of this evaluation by comparing the statistic variables.



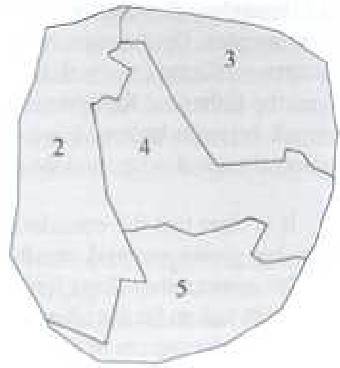
	707.0	737.0	757.0	777.0	797.0
Maximum	0.4188	0.3837	0.3786	0.3743	0.3719
Minimum	0.1822	0.1988	0.2166	0.2303	0.2379
Difference	0.2366	0.1849	0.1620	0.1440	0.1340
Mean Deviation	0.0612	0.0422	0.0370	0.0329	0.0306
Mean	0.2816	0.3167	0.3199	0.3221	0.3234

Weight

Figure 5.13: Result for increasing weight value for criteria two or Fractal Dimension

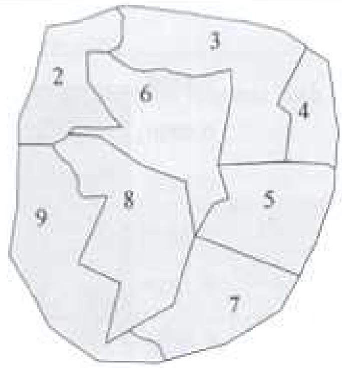
Considering Euclidean Measure only

	AREA	*PERIMETER*	*ICI*
2	2083104.66	7673.25288	0.4472
3	1675388.64	5777.51303	0.4472
4	1800895.96	7532.07689	0.4472
5	1485328.86	6699.33352	0.4472
Mean Deviation	***180820.78***	***682.1208***	***0.0000***
Mean	***1761179.53***	***6920.5441***	***0.4472***



Considering Area Size only

	AREA	*PERIMETER*	*ICI*
2	1130688	5785.41771	0.2033
3	635196.2	4049.68876	0.2033
4	374175.9	2748.23546	0.3338
5	1033065	5684.69855	0.0997
6	1151812	5398.02664	0.2033
7	1145671	6085.91901	0.0997
8	799795.7	3697.89148	0.4642
9	774314.6	3862.1412	0.4642
Mean Deviation	***234719.16***	***1074.5131***	***0.1213***
Mean	***880589.76***	***4664.0024***	***0.2589***



Considering Fractal Dimension only

	AREA	*PERIMETER*	*ICI*
2	508668.3	3738.41014	0.3086
3	529576.9	3023.68786	0.3723
4	734108	3853.49572	0.3086
5	314743.9	2323.98352	0.3086
6	674855.9	4042.4062	0.3086
7	630279.9	4168.58342	0.3086
8	1136107	5188.92926	0.3086
9	738567.5	4007.25613	0.3086
10	1009470	4794.77727	0.3086
11	768341.1	3570.62467	0.3086
Mean Deviation	***172846.81***	***569.1750***	***0.0115***
Mean	***704471.81***	***3871.2154***	***0.3150***

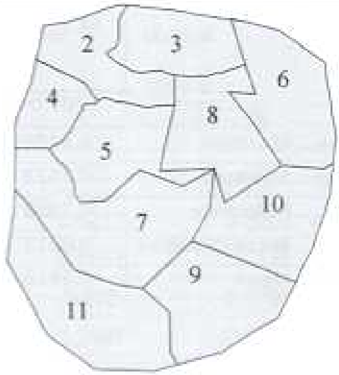


Figure 5.14: Result for considering one criteria only

The gradual increase of the Mean values and decrease of Mean Deviation value shows that when the weight value of criteria two, the Fractal Dimension is increasing gradually, the output result give an improving result. Finally, another evaluation is done to seek the performance of the Enhanced Redistricting Algorithms if only one criterion is used at each time. At first, the research only considers the Area Size, then the research considers the Fractal Dimension only and finally, the research considers the Euclidean Measure only to produce a set of different district plans by using the prototype. The output result is studied based on the statistic in Figure 5.13

With only Euclidean Measure based on Area-perimeter ratio, some of the district boundary complexities are high and the shape is odd. There are only four districts showing that the district plan is not fulfilled the requirement of the redistricting goals although the Mean values for the ICI is high and the Mean Deviation values is zero. Then, with considering the area size only, the district complexity is high because the Mean values for ICI is low and the Mean Deviation is in large value. On the other hand, with considering the Fractal Dimension only, the Mean value for the Area Size is not convincing. Subsequently, with one criterion or single measurement alone, the output result is not representative according to the result and analysis. In brief, it is clear that with the consideration of three criteria on Fractal Dimension, Area Size and Euclidean Measures, the output district plan is optimal compact and fulfills the redistricting goals.

5.2.4 Decision Indexing

At each stage, the decision indexing or namely the Integrated Compactness Indexing (ICI) defines qualified districts. Therefore, the research needs to evaluate the performance of the overall prototype for different values of the decision indexing at each stage in the Dynamic Programming Module to fully understand the performances of the FMCDM module. Thus, Figure 5.15 and Figure 5.16 show the performance of the Enhanced Redistricting Algorithm at different decision indexing values at 0.25, 0.3, 0.35, 0.4, 0.5 and 0.6. Figure 5.15 shows the Mean value of the ICI in the output value and Figure 5.16 shows the Mean Deviation of the ICI for the relative measure at the district plan. The performance of the algorithm starts to work consistently when the decision indexing is greater than 0.4. It is noticed from the chart in Figure 5.16 that the output plans with stable ICI values. Besides, the Mean Deviation dropped exponentially shows that the algorithm performs better when the decision indexing is greater than 0.4.

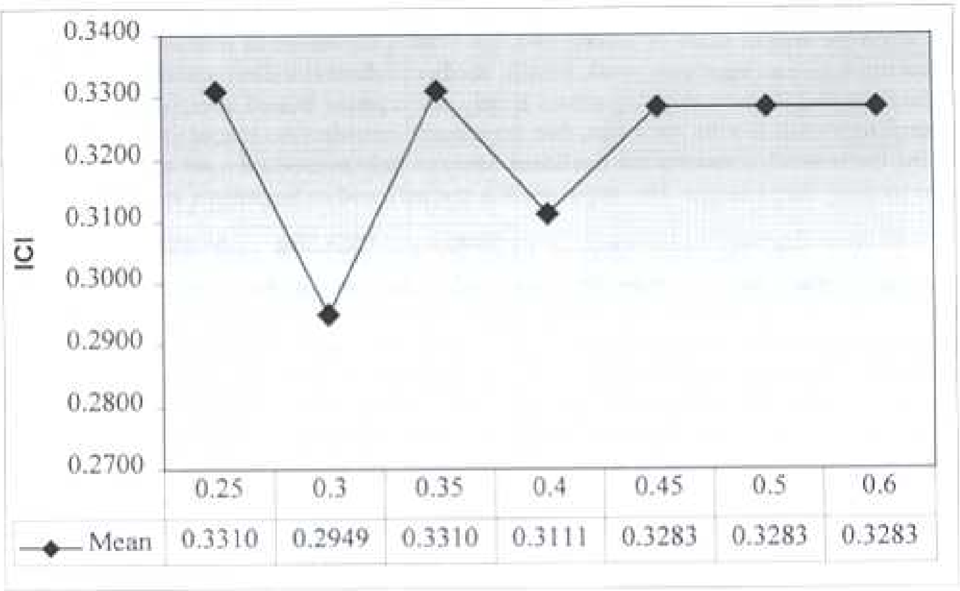


Figure 5.15: Mean value for ICI of the output result at different decision indexing.

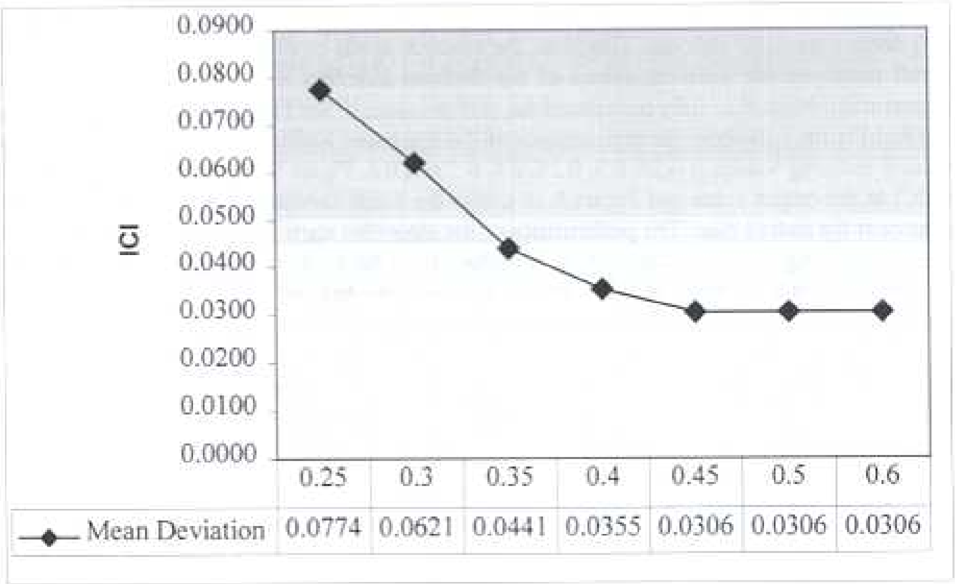


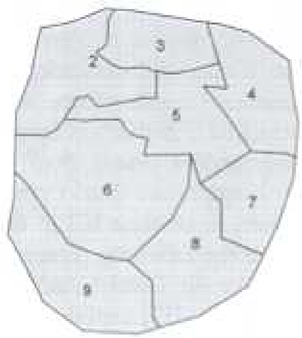
Figure 5.16: Mean Deviation for ICI of the output result at different decision indexing.

5.3 Overall Performance Analysis

This section emphasizes on the results and analysis for the overall performances of the redistricting algorithm based on the case study. The application dependent criteria is based the District Area Size and slope data calculated from contour line and the application independent criteria in based the indexing from two selected compactness measurements like Euclidean Measure and Fractal Dimension. The estimation of the Integrated Compactness Indexing (ICI) is using both the application dependent and application independent criteria from the attribute databases as its input. Thus, the discussion in this section compares two district plans generated with and without the integration of the FMCDM component to produce ICI in the enhanced algorithm. The district plan generated from the algorithm uses the following parameter: triangles size = 5,000; ideal district size = 1,000,000; weighing vector for District Area Size, Fractal Dimension, Euclidean Measure on area-perimeter ration = triplet 717; confidence level = 0.5; attitude to risk =0.5. All these variables are entered during the initiation process of the prototype (Refer Figure 5.1and Figure 5.5).

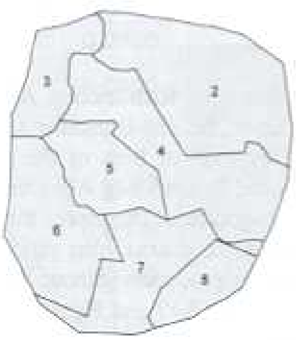
Figure 5.17 and Figure 5.18 show comparisons results on the overall performances. For the sake of evaluation between the results, this research uses two selected compactness measurements like Euclidean Measure based on Area-perimeter ratio (EM) and the Fractal Dimension (FD) to calculate the compactness indexing for each of the district plan. In comparing the district compactness, the smaller the Euclidean Measure value and the closer the Fractal Dimension value to 1.0000, the more compact is the district. Statistics methods such as Maximum, Minimum, Difference, Mean, and Mean Deviation are used to analyze the performances of both algorithms.

The results from this analysis show that the Enhanced Redistricting Algorithm can produce more compact district plan because its Mean value of the EM is smaller and the FD is closer to 1.0000 (FD for circles). Besides, the Mean Deviation of both of the compactness indexing is lower in the district plan generated by the Enhanced Redistricting Algorithm compared to the district plan produced by without the enhanced algorithm. Therefore, the Difference between the Maximum and Minimum value for the indices in district plan produced by the algorithm is smaller than the Difference of the district plan produced without the algorithm. Besides, the district plans and their statistical result in the Figure 5.17 and Figure 5.18 show the district plan without the Enhanced Redistricting Algorithm is odd or bizarre. For instance, the district number 7 has an area that only connected with a point. Furthermore, the district area size does not give consistent value on the Area size. For example, the District No. 2 give an Area size that is about 70 percent larger than the supposed area size (ideal Area size is 1,000,000). Besides, the District No. 8 has area size that is about 60 percent smaller than the supposed value. If the research compares the district perimeter, the district plan produced without the algorithm gives the range from around 2000 to 7000 but the district plan produced with the algorithm give the range around 3000 to 5000. This shows that the district boundary for the district plan without the algorithm is much more complex or less compact. In overall, the improvement of the redistricting algorithm based on Euclidean Measure for the district plan of the Enhanced Redistricting algorithm has the improvement for Mean at around 15.5 percent meanwhile the improvement for Mean Deviation is up to 41.8 percent.



No.	AREA	PERI-METER	SLOPE	EM	FD
2	854779	4819	31.617	27.1681	0.9977
3	529576	3023	24.743	17.2641	0.9966
4	737343	3938	33.861	21.0385	1.0109
5	859656	5627	24.743	36.8357	1.0101
6	1550218	5097	14.590	16.7595	1.0019
7	564867	3059	26.952	16.5677	0.9981
8	938806	4628	22.204	22.8210	1.0125
9	1009469	4794	44.248	22.7742	1.0133
Maximum				**36.8357**	**1.0133**
Minimum				**16.5677**	**0.9966**
Difference				**20.2680**	**0.0167**
Mean Deviation				**4.6871**	**0.0063**
Mean				**22.0087**	**1.0062**

Figure 5.17: District plan produced with the Enhanced Redistricting Algorithm



No	AREA	PERI-METER	SLOPE	EM	FD
2	1752461	6051	33.861	20.8974	1.0018
3	554490	3590	31.617	23.2555	1.0097
4	1468443	7164	31.617	34.9536	1.0041
5	694869	3680	31.617	19.4990	0.9979
6	1134549	4788	43.401	20.2088	1.0186
7	990493	6704	44.248	45.3785	1.0240
8	449410	2859	0.000	18.1919	1.0036
Maximum				**45.3785**	**1.0240**
Minimum				**18.1919**	**0.9979**
Difference				**27.1866**	**0.0261**
Mean Deviation				**8.0635**	**0.0076**
Mean				**26.0550**	**1.0085**

Figure 5.18: District plan produced without Enhanced Redistricting Algorithm

5.4 Conclusion

This chapter analyzes the result and analysis of the Enhanced Redistricting Algorithm of the prototype through statistical method. The prototype not only enables users to select their necessary input data and output name, but also triangle size and ideal district size for the redistricting process. Furthermore, DM can make decision by assigning weighing vectors for each considered criteria, confidence level, and attitude to risk. Therefore, the algorithm has the ability to consider multiple criteria with respective importance to ensure the optimality in compactness when producing a district plan. The algorithm has good improvement compared to redistricting algorithm without the enhanced component and it shows that the result with up to 41 percent of improvement in the Mean Deviation for the relative ICI and also 15 percent of improvement on the Mean ICI value. The analysis shows triangle size which is 0.005 percent over the ideal district size is the optimal value to produce the district plan. Besides, the analysis concludes that the shape of the license boundary will affect the algorithm. Rectangular and circle shape of the license boundary will produce more compact district plan compare to other shape. In respect of the Fuzzy Multiple Criteria Decision Making in the integration, several conclusions are made. Firstly, definition of the linguistic term for the shape optimal rules based on circles, octagon and hexagon are better than rectangles. Secondly, observation parameters can adequately reflect the uncertainty and impressions associated with the DM's subjective judgement in human thinking. With these variables, user may classify into optimistic, moderate, or pessimistic. Thirdly, weighing vector used in the algorithm is extremely useful to initiate the importance of each criterion. Lastly, the decision indexing or the Integrated Compactness Indexing for each stage is analyzed and the value of 0.5 if the optimal value to generate optimal district plan. Finally, the overall performance analysis indicates that the algorithm for shape-based redistricting is able to draw optimal district plan. In conclusion, the algorithm really shows ability to give better performance to draw or redraw district plan with optimal compactness.

CHAPTER 6: SUMMARY AND FUTURE WORK

The main focus of this chapter is to discuss the findings on the shape-based redistricting algorithm developed for the integration of multiple compactness measurements with Fuzzy Multiple Criteria Decision Making after it had been conceptually designed, implemented, and evaluated. The main accomplishment of the research is the successful development of an algorithm to incorporate multiple compactness measurement into redistricting technique. The enhanced algorithm used a new compactness indexing to evaluate the compactness and continuity of a district plan. The new compactness indexing was called Integrated Compactness Indexing because it was integrated with Fuzzy Multiple Criteria Decision Making for multiple compactness measurement. Besides, the index was also considering an application dependent criterion for redistricting goals. The index was used for a shape-based redistricting based on Dynamic Programming to generate optimal compact district. The achievement showed the contribution of the developed redistricting approach in improving redistricting technique.

Besides, it achieved the main objectives that are stated in Chapter One of this dissertation. Though the redistricting algorithm had its limitations those could not be resolved, a list of future works was discussed in the following sections to enhance the algorithm in future.

6.1 Research Findings and Contribution

The achievement of the predefined objectives has proven the feasibility of the developed shape-redistricting algorithm to produce an optimal compact district plan. The overall thesis has presented a complete procedure, design, implementation, development, operation and performance of an integrated compactness indexing with Fuzzy Multiple Criteria Decision Making for shape-based redistricting. Hence, it proves that the research framework has produced the desired output.

The research has studied current algorithms and methods on redistricting to identify strengths and weaknesses of the current algorithms regarding to their performance, accuracy and factors considered. In brief, although there were tremendous efforts in improving the redistricting algorithm, 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 aspect 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 involved multiple criteria. The complexity of the geographical-related problem has made the redistricting process to generate odd or not-practical shape for district plan. Therefore, all the problems found have stimulated the idea of considering multiple criteria especially compactness measurement in decision making of a geographical aspects or namely shape-based redistricting algorithm.

The research has also identified several important requirements for shape-based redistricting algorithm. Firstly, the algorithm should use multiple compactness measurement method in order to gather the strengths of particular method and to reduce or minimize its weaknesses or lacking. Thus, multiple criteria that including both of the application dependent and application independent factors like compactness should be used to produce an integrated shape compactness indexing that was more descriptive and able to incorporate with natural feelings of decision makers. Natural feelings here included their confidence and their attitude to risk. The consideration of the multiple criteria was able to cope with vagueness or fuzziness. After that, the shape-based redistricting should use the indexing to assist the generation of optimal compact district plan. Besides, it should able to consider the restricted boundary like river, political boundary and any others during the redistricting process. Lastly, the new algorithm should work and perform in an environment (software or hardware interface) that is able to manage the spatial and non-spatial data, their relationship and dependency in order to simplify the whole shape-based redistricting process.

The research has successfully designed and developed a shape-based redistricting algorithm. The use of a FMCDM method called Fuzzy-AHP was integrated with multiple criteria including the compactness measurement to produce optimal compact district plan. The algorithm has defined the procedures for knowledge acquisition, initiation, multiple criteria decision making and drawing of the district plan. An Integrated Compactness Indexing has been successfully calculated and used into redistricting technique. This compactness indexing was more descriptive and able to incorporate natural feelings of DM with a set of observation parameter on the confidence level and attitude to risk. The Fuzzy AHP approach allowed the integration and consideration of both Application Dependent and Application Independent factors in the redistricting application. 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. Besides, the shape-based redistricting technique was able to cope with fuzziness with the use of the triangular fuzzy number to define the decision matrix and weight vectors for each criterion.

Later, this research found and proved the applicability and effectiveness of the shape-based redistricting algorithm by using a prototype. This prototype allowed flexible demonstration, analysis and visualization of the output. The prototype demonstrated the application of redistricting algorithm developed and managed to demonstrate the concept of the integration of Fuzzy-AHP in compactness measurement indexing for the shape-based redistricting algorithm. The overall performance of the developed redistricting algorithm was evaluated under different circumstances and it clearly showed its advantages. These advantages included (a) better modeling of the uncertainty and imprecision associated with the fuzzy triangular number, (b) Cognitively less demanding on the district planner, and (c) adequate reflection of the district planner's attitude towards risk and their degrees of confidence in their subjective assessment through the analysis and result on different sets of α and λ values in Section 5.2. Real experience in applying the approach in selecting the most appropriate district plan by the algorithm implemented in the prototype had reinforced these findings.

Lastly, the research has found that GIS as a useful tool to fully support the functionality of shape-based redistricting technique, nevertheless it was GIS's nature to relate the map databases and attributes databases in order to perform redistricting data manipulation, analysis and visualization. Therefore, the research achieved its objectives and able to prove the result of the integration in all aspects. In summary, the research findings were able to contribute to all DM when conducting the shape-based redistricting.

6.2 Limitations of the Redistricting Algorithm

The developed redistricting algorithm faced a few limitations during building stage. The Redistricting prototype was built only for two-dimensional environment. Consequently, it did not permit for three-dimensional redistricting applications. Besides, the integration of multiple compactness measurement methods should able to gather the strengths of particular method and at the same time to reduce or minimize the weaknesses or lacks of other method. However, one must able to manipulate the weight vector that represents the importance of a particular measurement in order to accomplishes it. Therefore, a group of expert needed to have extensive knowledge on manipulation of these weight vectors. These weight vectors would then influence the integrated compactness indexing, which would affect the shape compactness of the district as well.

The working time of redistricting prototype was not satisfactory because of the computer system used to operate the prototype. The map databases were large and consumed a large hard disk space and higher memory power. Applications that used map databases need a lot processing power, as many layers of database needed to be loaded at once for analysis.

6.3 Future Work

After reviewing the findings, there are a number of areas of future works for the redistricting algorithm. First, extend the analysis and modeling of more Application Dependent and Application Independent information sources especially on their relationships. The current system does not do any detail analysis of these factors due the lack of data and complexity that cause the research to model and analyze them separately.

Second, study the relationships (correlation) between various shape compactness measurements. The detail analysis on their relationships may help their consideration in the algorithm to enhance the result of the redistricting technique. Presently, researches are mutually exclusive and no detail study has been done to analyze the their relationships. It needs to look for an algorithm to compensate the relationships to produce better district map.

Thirdly, integrate with raster map. Currently, the maps used are in vector format. It would be another challenge to the redistricting algorithm if other format of maps can be used such as raster map.

6.4 Concluding Remarks

The research has been completed successfully despite all the constraints being posed to. The most profound accomplishment among all is the success in creating a new compactness measurement indexing called Integrated Compactness Indexing and incorporating it into the redistricting algorithm to enhance its result. This domain knowledge will certainly useful in the knowledge theory. Though the information sources are limited, it is enough to prove the applicability of the developed redistricting algorithm to incorporate shape compactness information into redistricting technique for optimization of the compact district plan. The algorithm is simple and yet efficient as compared to other more complicated methods. Though the redistricting algorithm was tested under a few constraints especially on the Fuzzy Multiple Criteria Decision Making component, however the results is satisfactory and comparatively better than other traditional approaches as well. The mathematical framework for obtain the Integrated Compactness Indexing can be applied also for other redistricting applications and system. As a conclusion, further enhancement of the redistricting algorithm and prototype can be achieved without constraints when more sources of information are available and better technology can be obtained. Besides, future attempt can be done to apply or integrate more criteria to enhance the overall performance of the developed redistricting algorithm.

BIBLIOGRAPHY

Agnew, John (1994); *Geography and Regional Science Program: Geographic Approaches to Democratization*; National Science Foundation; Washington.

Altman, Micah (1997a); *What are the Judicially Manageable Standards for Redistricting? Evidence from History*; California Institute of Technology; Pasadena, California.

Altman, Micah (1997b); *Is Automation the Answer? The Computational Complexity of Automated Redistricting*; California Institute of Technology; Pasadena, California.

Altman, Micah (1998); *Redistricting Principles and Democratic Representation*; Ph.D. Thesis; California Institute of Technology; Pasadena, California.

Balas, Egon; Padberg, Manfred W. (1965); "Set partitioning: A Survey"; *SIAM*; p.p.710-760.

Baoding Liu, (1999); "Introduction to Mathematics Programming"; *Uncertain Programming*; John Wiley & Sons, Inc. United States of America.

Blevins, J (1999);*Political Redistricting:* Operations Research References; URL:http://ww.mai.lie.se/~jible/#Political Redistricting

Bourke Paul (1993); *Fractal Dimension Calculator User Manual*; URL:http://www.swin.edu.au/astronomy/pbourke/fractals/fracdim/

Brian A. Cromwell & William T. Bogart (2000); "How Much Is a Neighborhood School Worth?"; *Journal of Urban Economics 47*; p.p.280-305.

Brochmann, Harold (1999); *Fractal Geometry* URL:http://www.saltspring.com/brochmann/Fractals/fractal-0.00.html

Browdy, Michelle H. (1990). "Simulated Annealing: An Improved Computer Model for Political Redistricting." *Yale Law & Policy Review 8*: p.p.163-179.

Bowdren , Peter (1997); *Determination of the Fractal Dimension;* URL:http://life.csu.edu.au/fractop/doc/notes/first.htm

Cain, Bruce (1984); *The Reapportionment Puzzle;* California: University of California Press.

Carmen Cirincione, Thomas A.D, Timothy G.O. (2000); "Assessing South Carolina's 1990s Congressional Redistricting"; *Political Geography, 19*; p.p. 189-211; Pegamon.

Chen Shyi-Ming (1996); "A new Method for Evaluating Weapon System Using Fuzzy Set Theory"; *IEEE Transaction on Systems, Man and Cybernetics Part A: System and Human*; Vol 26, p.p. 493-496.

Clark, Michael, Tobin Crain, Shane Groves, and Robert White (1999); *Gis Application For Redistricting In The State Of Texas*; Texas Legislative Council (TLC).

Cowen, Lenore J. & Rachel T. Manley (1998); *Approximation Algorithms*; Department of Mathematical Sciences; Johns Hopkins University; URL:http://www.cse.buffalo.edu/~rtmanley/perim.html

DeCola, L., & N. S. Lam (1993); *Introduction to Fractals in Geography*; Fractals in Geography; Englewood Cliffs, N.J.: Prentice Hall.

Deng H.(1999); "Multicriteria Analysis with Fuzzy Pairwise Comparison"; *International Journal of Approximate Reasoning* 21(3): p.p. 215-231.

Dominique Bérubé & Michel Jébrak (1999); "High Precision Boundary Fractal Analysis for Shape Characterization"; *Computers & Geosciences, 25* ; p.p.1059 –1071.

Duonne M Erasmus (1996); *Practical GIS for Managing Forestry, Operations and Planning*; Mondi Limited, South Africa.

Fred Hejazi & Monica Dombrowski (1997); *Automated Redistricting System for Law Enforcement*; URL:http://www.dande.com/digital/digital.htm

Fullér ,Robert (1996); *Lecture Notes : Fuzzy Decision Making*; URL:http://www.abo.fi/~rfuller/fdm.html

Fuller Robert and Carlsson Christer (1996); "Fuzzy Multiple Criteria Decision Making: Recent Developments"; *Fuzzy Sets and Systems 78*; p.p. 139-153.

Galaune Alfredas & Mozgeris Gintautas (1998); "Geographic Information Systems in Lithuanian Forestry: An Arc/Info And ArcView Based Solution"; *13th ESRI European User Conference*; Firenze Italy.

George J.Klir, Tina A. Folger (1992); *Fuzzy Sets, Uncertainty and Information*; Prentice Hall International Editions; Singapore.

Greenberg, Harvey J. (1998); *Greedy Algorithms*; University of Colorado at Denver.

Greenberg, Harvey J. (1999); *Mathematics Programming Glossary* ; URL:http://www.cudenver.edu/~hgreenbe/glossary/glossary.html

Handley, Lisa & Maley, Michael (1997); *Boundary Delimitation Index*; URL:http://www.aceproject.org/main/english/bd/default.htm

Hayes, Bayes (1997); *Computing Science, Machine Politics*; URL:http://www.sigmaxi.orgamsci/Issues/Comsci96/compsci91-11-source.html

Hearne W. Walter (1999); *Methods to Madness: Alternative Schemes for Single-Member Redistricting*; The Center for Voting and Democracy.

Hepu Deng (1999); "Multicriteria Analysis with Fuzzy Pairwise Comparison"; *Fuzzy Systems Conference Proceedings, 1999. FUZZ-IEEE '99*; 1999 IEEE International, Vol. 2; p.p.726 –731.

Herrera, F. and Verdegay, J.L. (1995); *Fuzzy Sets and Operations Research Perspectives*; Technical Report; Department of Computer Science and Artificial Intelligent; Granada, Spain.

Hershberger J. & Subhash Suri (1991);"Finding Tailored Partitions"; *12 Journal of Algorithm 431*.

Hoover, Aaron (1997); *Can We Give Children The Education They Need?* ; Wilmington Star-News, Inc.; Southeastern North Carolina.

Jacob Jen-Gwo Chen and Zesheng He (1997); "Using Analytic Hierarchy Process and Fuzzy Set Theory to Rate and Rank the Disability"; *Fuzzy Sets and System, 88*; p.p.1-22.

Jones, Lillie (1998); *Redistricting Guilford County Schools*; Guilford County Board of Education; Greensboro News & Record, Inc.

Kashani, Reza (1998); "Fuzzy Logic Decision Making"; *DEICON Dynamics & Control*; URL:http://www.deicon.com/case_s_html/dec_make/index.htm

Klamroth Kathrin and Wiecek, Margerat M. (1999); *Dynamic Programming Approaches to the Multiple Criteria Knapsack Problem*; Clemson University; Clemson, SC, USA.

Knight , Jeffrey L. (1997); *GIS Based Compactness Measurement Using Fractal Analysis*; George Mason University Fairfax, Virginia.

Kenneth C.Laudon & Jane P.Laudon (1996); *Management Information System*; Prentice Hall International, Inc; New Jersey.

Laurini R. & Thompsom D. (1992); *Fundamentals of Spatial Information System*; Academic Press; London.

Liittschwager, John M. (1973); "The Iowa Redistricting System. In Democratic Representation and Apportionment: Quantitative Methods, Measures, and Criteria"; *New York: Annals of the New York Academy of Sciences*; p.p. 221-234.

Matthew Warfel (1998); Fractal Page; URL: http://www.cee.cornell.edu/~mdw/fractal.html

Mehrotra, A., Johnson, E. L. & Nem Haouser, G. L. (1998); "An Optimization Based Heuristic For Political Redistricting"; *Management Science, 44(8)*; p.p.1100-1114.

McNeill, Daniel & Feiberger, Paul (1993); *Fuzzy Logic*; Simon & Schuster; New York.

Modell, Martin (1987); *A Professional Guide to System Analysis*; McGraw-Hill Software Engineering Series; United States of America.

Mon, Don-Lin (1995); *Evaluating Weapon System Using Fuzzy Analytical Hierarchy Process Based On Entropy Weight*; IEEE ; p.p. 598-591.

Muller. J. C. (1986); "Fractal Dimension and Inconsistencies in Cartographic Line Representations"; *The Cartographic Journal 23*; p.p.123-130.

Newsom J. (1997); *In-Depth: Redistricting Our Schools: Some Facts about School Redistricting*; Depot News; Piedmont Triad, North Carolina.

Nelson G.J (1998); *Context Aware and Location System*; A Dissertation submitted for the degree of Doctor of Philosophy in the University of Cambridge.

Nowicki A, et. al. (1997); *ArcView in the Polish Forestry* ; ESRI.

Ngu Meng Siong (1998); *GIS Based Application for Redistricting the Election Boundaries* (Unpublished); A final year research for Bachelor Degree study of Faculty of Information Technology, Universiti Malaysia Sarawak (UNIMAS).

Ohnishi, Shi-ichi and Imai, Hideyuki (1998); *Evaluation for Stability of Fuzzy Decision Making Using a Sensitivity Analysis*; IEEE.

Pamela McCauley Bell and Heng Wang (1997); "Fuzzy Linear Regression Model for Assessing Risks for Cumulative Trauma Disorders"; *Fuzzy Sets and Systems, 92*; p.p.317-340.

Paul Bourke (1993); *Fractal Dimension Calculator User Manual*; URL:http://www.swin.edu.au/astronomy/pbourke/fractals/fracdim/

Rawlin, Gregory J. E. (1998); *Genetic Algorithms*; Indiana University Computer Science; URL:http://grouchy.cs.indiana.edu/l/www/research/index.html

Ravi, V. and Reddy, P.J. (1999); "Ranking of Indian Coals via Fuzzy Multi Attribute Decision Making"; *Fuzzy Sets and Systems, 103*; p.p. 369-377.

Render, Barry et. al. (1997); "Chapter 17: Dynamic Programming (pg. 780-800)"; *Quantitative Analysis for Management*; International Edition; Prentice Hall Inc., London.

Robert Cheetham & Kevin Switala (1997); *Police at Play: What if Redistricting Plans were as addictive as Solitaire(tm)?* ; Philadelphia Police Department , Philadelphia.

Ron Kwok et. al. (1998); *Improving Group Decision Making : A fuzzy GSS Approach*; Dept of Information System, City University of Hong Kong; Tat Chee Avenue, Kowloon, Hong Kong.

Ruong, Xu (2000); "Fuzzy Least-squares Priority Method in the Analytic Hierarchy Process"; *Fuzzy Sets and System 112*; p.p. 395-404.

Sedgewick, Robert (1989); Algorithms (Second Edition); Addison-Wesley Publishing Company, Inc. Canada.

Sniedovich, Moshe (1992); *Dynamic Programming;* Marcel Dekker, Inc.;United States of America.

Sniedovich, Moshe (2000); *WORMS Virtual Encyclopedia*; Department of Mathematics and Statistics; URL:http://www.worms.ms.unimelb.edu.au/encyclopedia/frame.html.

Shiode, Narushige (1998); *Towards Application Of GIS For Visualizing Fractal Phenomena*; Center For Advanced Spatial Analysis; University College London.

Shooman Martin L. (1983) *Software Engineering: Design, reliability and Management.* McGraw-Hill international Editions, Computer Science Series.

Skiena, Steven , (1996); *Analysis of Algorithms*; Department of Computer Science; SUNY Stony Brook; URL:http://www.cs.sunysb.edu/~algorith/lectures-good/index_P.html

Skiena, Steven, (1997); *The Algorithm Design Manual*; Department of Computer Science State University of New York, Springer-Verlag, New York; URL:http://evo.apm.tuwien.ac.at/AlgDesignManual/BOOK/BOOK5

Terral, Rolanda (1996); *Redistricting and the African-American Community*; URL:http://www.bivote.org/ncbvp/background/redistricting

Triantaphyllou Evangelos and Stuart H. Mann (1995); "Using The Analytic Hierarchy Process For Decision Making In Engineering Applications: Some Challenges"; *International Journal of Industrial Engineering: Applications and Practice*; Vol. 2, No. 1, p.p. 35-44.

Vickrey, William (1961); "On the Prevention of Gerrymandering"; *Political Science Quarterly 76*; p.p. 105-109.

Wang, Li-Xin (1997); *A course in Fuzzy Systems and Control (International Edition)*; Prentice-Hall PTR; United States of America.

Weisstein Eric & Stephen Wolfram (1999); *MathWorld;* CRC Press; URL:http://mathworld.wolfram.com/BrelazsHeuristicAlgorithm.html

Yeh C.H., and Deng H.(1997); "An Algorithm for Fuzzy Multi-Criteria Decision Making"; *Proceedings of the IEEE First International Conference on Intelligent Processing Systems (ICIPS'97)*; p.p. 1564-1568.

Yeh C.H, Deng H and Pan H (1999); "Multi-Criteria Analysis For Dredger Dispatching Under Uncertainty"; *Journal of the Operational Research Society 50*; p.p. 35-43.

Zimmermann H.J. and Sebastian H.J. (1995); "Intelligent System Design Support by Fuzzy Multi-Criteria Decision Making and/or Evolutionary Algorithms"; *Fuzzy Systems, International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium., Proceedings of 1995 IEEE International Conference*; Vol. 1, p.p. 367 -374.

APPENDICES A

A.1 Altman's Axioms Evaluation on Compactness Measurement

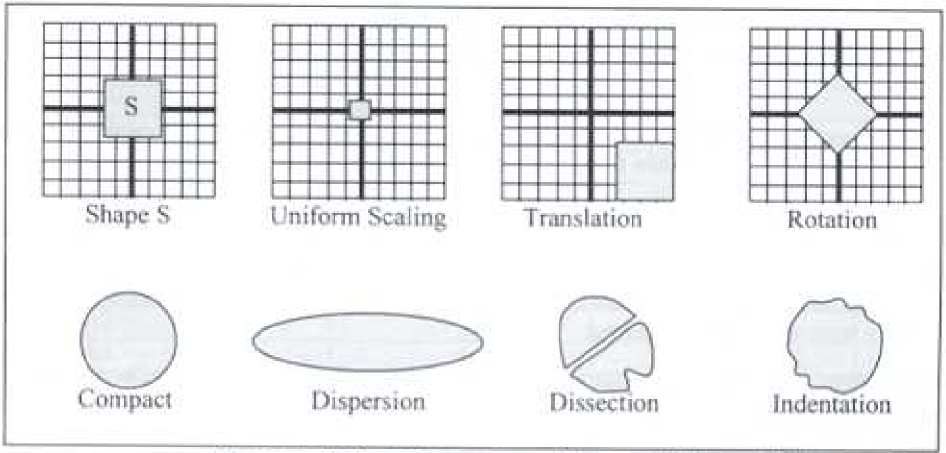


Figure A.1: Transformation of shapes. (Atlman, 1998)

A.1.1 Definition

Altman (1998) uses an axiomatic analysis to join formal measures of compactness to analyze the consistency of these different compactness criteria. First, the research will need some definitions:

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 $Area\ (s_i \cap s_j) = Perimeter\ (s_i \cap s_j) = 0, \forall i \neq j$ (refer Figure 3.2). Let $P{:}S \to \Re$, be the length of the perimeter of the shape, and let $A{:}S \to \Re$, be the area of the shape. Let a compactness measure C, be a function $C{:}S \to \Re$.

Using these definitions, the research can now formally define what it means for a compact measure to capture shape (Figure A.1):

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.

A compactness measure must not violate any of these three principles. It would be strange indeed if the research could change a district's shape simply by uniformly scaling, rotating, or moving the map upon which it is drawn.

In the next three principles, Altman(1998) captured the concepts of dispersion, dissection, and indentation. First, compactness measures that claim to capture dispersion are usually based on the ratio of a shape's perimeter to its area. These measures work well for convex shapes, but can confuse indentation and dispersion for nonconvex shapes (Figure A.2).

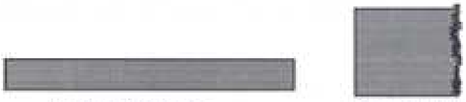


Figure A.2: The Perimeter/Area ratio fails to capture dispersion. The P/A of the figure on the left is less than that of the one on the right.

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. The research can 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:

5. *Minimal dissection*: Any two shapes with identical convex hulls, S_1 has a strictly smaller area, then S_1 should be judged 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 less compact.

In addition to capturing recognized methods of manipulation, any compactness measure that satisfies any of these axioms will have two other nice properties.
Property 1 (Continuity): For any given perimeter or area, the maximally compact shape is contiguous. If C satisfies axiom 6, this is true for any given convex hull as well.
***Property* 2**: A circle is the most compact district.

A.1.2 Result and Proof

The principles in section 2.3.1 set bounds on a reasonable compactness standard — if a compactness measure contradicts three of principles of compactness (or any of the shape principles), the research should suspect it of measuring something other than geographic compactness. Table A.1 summarizes the results of applying these axioms. The details and descriptions for each of the compactness measurement in Table 3.1 can be found in Section 2.2.

Altman (1998) draws some conclusions from his axioms examinations. First, although most of the compactness measures meet the first three axioms, eight measures violate, in their standard form, the basic axioms for measuring shape. Then, three measures, LW5 , OS2 and PA5 , clearly violate these axioms, and so should be rejected. Besides, five others, the remaining PA measures, in their current form, violate axiom 1, but they are safe to be used. However, the research needs to be careful to measure the boundaries of all districts with the same precision when the research uses these measures. Moreover, 12 compactness measures (LW1–LW7, A4–A7, OS1 and OS4) violate all last 3 axioms of compactness. In other words, they do not exactly deal with geographically aspect of redistricting.

Table A.1:Violations of the measurement axioms by compactness are marked by a 'V' in the cells. (Altman, 1998).

Measure	*Axiom 1*	*Axiom 2*	*Axiom 3*	*Axiom 4*	*Axiom 5*	*Axiom 6*
LW1				V	V	V
LW2				V	V	V
LW3				V	V	V
LW4				V	V	V
LW5		V		V	V	V
LW6				V	V	V
LW7				V	V	V
A1				V		V
A2				V		V
A3				V		V
A4				V	V	V
A5				V	V	V
A6				V	V	V
A7				V	V	V
PA1	V*			V	V	
PA2	V*			V	V	
PA3	V*			V	V	
PA4	V*			V	V	
PA5	V			V	V	
PA6	V*			V	V	
OS1				V	V	V
OS2		V		V	V	V
OS3				V		V
OS4				V	V	V

(* In these cases the measure is sensitive to the scale of the measuring unit used to measure the district boundaries, not to the scale of the map upon which the district boundaries are represented)

Altman proves of these results in each measure and the axioms that it violates. Measures OS3 and A1, A2 and A3 clearly satisfy axioms 1–3 and 5 thus meeting the axiomatic criteria. However, most of the other measurements violate at least one of the first three, or all of the latter three axioms, raising doubts as to their consistency.

Many of these indices violate at least one of the first three "shape" axioms:

- Measure PA5 violates axiom 1. Convex districts of exactly the same shape, but different sizes may be assigned different values. All of the perimeter/area measures, PA1 to PA6, are subject to a more indirect violation of scale invariance in practice, which has not been previously recognized. If districts have natural boundaries, these measures can be affected by precisely how the research measures district lines, for districts will seem to be less compact when seen on a map which has a fine scale than on a map with a larger scale. For comparisons to be consistent, the research must use the same precision to measure all district lines.
- For any finite number of sample points, chosen at fixed positions along the edge of the shape, OS2 violates axiom 2, because rotating a shape may change the choice of sample points, and hence the compactness measurement. Measure LW5, by its definition, fails axiom 2.

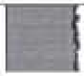
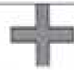
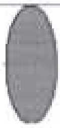
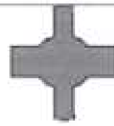
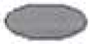
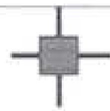
More Compact (Under Axiom 4)	Less Compact (Under Axiom 4)	Measures Not Classifying These Correctly
		PA 1 -PA 6
		LW 1 -LW 5
		LW 6 ,LW 7 , A 1 ,A 2 , OS 3, A 4 -A 7
		OS 1, OS 2
		OS 4 ,A 3

Figure A.3 Violations of axiom 4. Shapes on the left have the same area, but smaller convex hulls than those on the right (Altman, 1998).

Most compactness indices reflect at least one principle of shape manipulation, but not others. In most cases, these measures obviously satisfy one shape axiom, but violate others. Altman demonstrate these violations by producing shapes that are misclassify by particular measures.

- Measures A1, A2, A3, and OS3 clearly satisfy axiom 5, although they violate axioms 4 and 6 (Figure A.3).
- Measure OS1 violates axioms 4, 5, 6 (Figure A.3, Figure A.4).

A number of compactness indices violate all three principles:

- All measures listed violate axiom 4 (Figure A.2).
- All with the exceptions of PA1–PA6 and OS2 violate axiom 6, because changes in perimeter that do not affect convex hull, area and shape diameters are ignored.
- Measures LW1–LW7, A4–A7 and OS4 violate all three compactness axioms (Figure A.3, Figure A4).

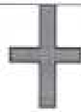
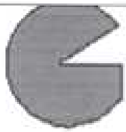
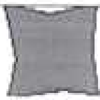
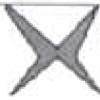
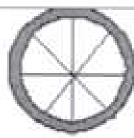
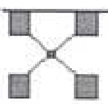
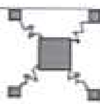
More Compact (Under Axioms 5,6)	Less Compact (Under Axioms 5,6)	Measures Not Classifying These Correctly
		LW 1 -LW 5
		LW 6 -LW 7, A 4 -A 6
		A 7
		OS 4
		OS 2
		OS 1

Figure A.4: Violations of axioms 5 and 6. Shapes on the left have the same convex hulls, greater area and a smaller perimeter/area ratio than those on the right(Altman, 1998).

A.2 Concept of Fuzzy Set Theory

A fuzzy number is a convex fuzzy set (Deng, 1999), characterized by a given interval of real numbers, each with a grade of membership between 0 and 1. Its membership function is piecewise continuous. A triplet can define a triangular fuzzy number (a_1, a_2, a_3). The membership function is defined as

$$\mu A(x) = \begin{cases} \dfrac{x - a_1}{a_2 - a_1}, a_1 \le x \le a_2, \\ \dfrac{a_3 - x}{a_3 - a_2}, a_2 \le x \le a_3, \\ 0 \quad , otherwise. \end{cases} \qquad (3.1)$$

where a_2 is the most possible value of fuzzy number A, and a_1 and a_3 are the lower and upper bounds respectively which is often used to illustrate the fuzziness of the data evaluated.

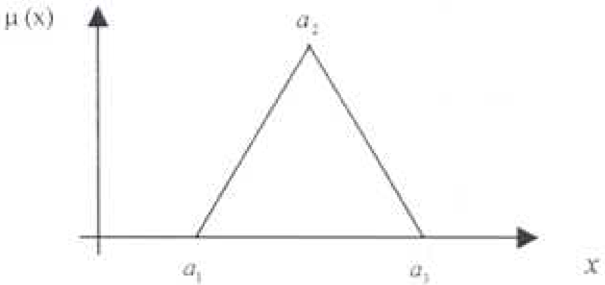


Figure A.5: Membership functions, μ(x) and linguitics terms for effectiveness (X)

Let $\overline{A}$ and $\overline{B}$ be to triangular fuzzy number, where

$$\overline{A} = (a_1, a_2, a_3), \qquad \overline{B} = (b_1, b_2, b_3)$$

The fuzzy number arithmetic operation can be summarized as follows:

(a) Addition: $\overline{A} + \overline{B} = (a_1, a_2, a_3) + (b_1, b_2, b_3) = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$

(b) Subtraction: $\overline{A} - \overline{B} = (a_1, a_2, a_3) - (b_1, b_2, b_3) = (a_1 - b_3, a_2 - b_2, a_3 - b_1)$

(c) Multiplication: $\overline{A} \times \overline{B} = (a_1, a_2, a_3) \times (b_1, b_2, b_3) = (a_1 \times b_1, a_2 \times b_2, a_3 \times b_3)$

(d) Division: $\overline{A} \div \overline{B} = (a_1, a_2, a_3) \div (b_1, b_2, b_3) = (a_1 \div b_3, a_2 \div b_2, a_3 \div b_1)$

(e) Inverse: $A^{-1} = (\dfrac{1}{a_3}, \dfrac{1}{a_2}, \dfrac{1}{a_1})$

A.3 Dynamic Programming
(Kenneth & Jane, 1996; Baoding, 1999; Sedgewick, 1989; Sniedovich, 1992)

Dynamic programming (DP) is a quantitative analysis technique that has been applied to large, complex problems that have sequences of decisions to be made. It is a mathematical programming technique, which is widely used in operations research. Comparing to other Mathematics Programming Methods such as Linear programming, Non-linear Programming, Goal Programming and Integer Programming, DP is effective in looking for the "Best or Optimal" way especially in Knapsack Problem. Dynamic programming divides problems into a number of decision stages; the outcome of decision at one stage affects the decision at each of the next stages.

Let us denote a multistage decision process by $[a, T(a,x)]$, where a is called state, $T(a,x)$ is called a transformation, and x is called decision vector. It is clear that the transformation function depends on the state a and decision vector x. We suppose that we have sufficient influence over the process so that at each stage we can choose a decision vector x from the allowable set S. Let x_i be the choice at the ith stage, then we have the following sequence,

$a_1=a_0$, **(an initiate state)**

$a_n=T(a_{n-1}, x_{n-1})$, $n = 2,3,\ldots$

We shall be concerned with processes in which the decision vectors x_i's are chosen so as to optimize a prescribed real-valued function of the state and decision vectors,

$R(a_1, a_2, \ldots; x_1, x_2, \ldots)$

which is called the criterion function, or return function. A decision is called optimal if it optimizes the criterion function.

In view of general nature of the criterion function R, the decisions x_n's are dependent upon the current state of the system as well as the past and future states and decisions. However, there are some criterion functions which have some special structures so that we can focus on the past history and current state of the process in searching the values of x_n's. With this assumption, we can represent the decision vector x_n at the stage n by the following form,

$x_n=x_n(a_1, a_2, \ldots, a_n; x_1, x_2, \ldots, x_n)$

which is called the policy function. Sometimes, we may have a simpler form of policy function, for example,

$x_n= x_n(a_n)$

for example , a function of current state.

Therefore, with DP, the optimal policy is characterized by Bellman's principle of optimality: *An optimal policy has the property that whatever the initial state and initial decision are, the remaining decision must constitute an optimal policy with regard to the state resulting from the first decision.*

A.4 Method for Fractal Analysis (Dominique & Michel, 1999)

Several methods are available for computation of the fractal dimension: structured walk, box counting, dilation and euclidean distance mapping (EDM).

A.4.1 Structure walk method

The structured walk is a vector-based method, which consists of walking around the perimeter of an object with a pair of compasses (divider) of a finite stride length. The perimeter of the outline is estimated from the number of steps needed to span the outline multiplied by the stride length. Changing the strike length, which is equivalent to changing the scale of observation, will produce another estimate of the perimeter. Plotting the log of perimeter against the log of the step length yields a linear relationship from which the fractal dimension (Fd) is derived by the relationship, $Fd = 1 - s$ where s is the slope of the plot and 1 represents the topological dimension of a line.

A.4.2 Box counting method

Box counting method is a matrix method. Averaging adjoining pixels to form a new image with coarser base elements progressively coarsens the representation of the object. The easier way for this method is to compute the dimension consists of placing a grid of finite box sizes on an outline, and counting the intersects between the boundary with the grid. Varying the box sizes in the grid produces a different number of intersections. Box counting has been widely used, mainly due to its versatility in the type of measurement (boundary, surface or lacunarity) and its ease of implementation.

A.4.3 'Sausage' method

The principle of 'sausage' method follows the same principles as the box counting method that it coarsens the initial image but in a more isotropic way. Circles of finite diameters are drawn around each point of the outline to form a ribbon, usually called covering or 'sausage'. The sausage method was mainly developed to ease the computerization of fractal analysis, thus avoiding the difficulties encountered with the structured walk method. Two implementations are presented as following.

A.4.3.1 Dilation method

Dilation is a computerized method to calculate the Fractal Dimension. Dilation is an image processing algorithm that adds a background pixel for every pixel in contact with the feature and erosion retrieves a feature pixel for every pixel that touches the background. Successful dilation operations on a one-pixel width outline produce a ribbon of a finite width. Dilation can also be obtained by the convolution of kernels of different diameter with outline. The resultant area for the kernel is divided by its diameter. The log on that result is then plotted against the log of the kernel diameter and the fractal dimension (Fd) is related to the slope of plot(s) through the relationship, $Fd = 1 - s$. The dilation process is straightforward and requires the simple application of filters available in image processing software.

A.4.3.2. Euclidean distance mapping

The EDM method is an image processing algorithm applied tp a black and white image, that results in a gray-scale image where each pixel is given a brightness value proportional to its distance to the nearest pixel of the outline of a feature. The feature then has a dark backbone that whitens up to the outline and gets darker again toward the background. The fractal dimension can be computed by plotting the log of perimeter against the log of the width of the covering circles. The fractal dimension (Fd) id derived from the slope of the plot(s) using the relationship, $Fd = 2 - s$ where 2 is the embedding space of the outline.

A.5 Prototype Coding

```
**********************************************************************************************
*           FILE NAME           : Start.aml            PURPOSE : startup file
**********************************************************************************************
&term 9999
display 9999 2 pos 10 10
&menu start1.menu &position &cc &screen &cc
&thread &delete &all
[close -all]

**********************************************************************************************
*           FILE NAME           : Message.menu        PURPOSE : To show message on status
**********************************************************************************************
7
 %msg1
%msg1 display .message 40
&return

**********************************************************************************************
*           FILE NAME           : Start.menu           PURPOSE : Create interface to get and assign necessary variables
**********************************************************************************************
7 Redistrict form
Redistrict - To redraw District line  by Integrated Compactness Index (ECI)
Input source %1
License source %2

River source %3

Name of Output      %4
District Size (sq m) %5  Triangular Size (sq m) %6
-----------------------------------------------------------------
Fuzzy Weight for Euclidean Measure:%7
Fuzzy Weight for Fractal Dimension:%8
Fuzzy Weight for Application Dependent Factor:%11
Degree of Confidence: %9  Attitude towards risk: %10
 %ok   %cancel   %help
%1 input .inshp 32 typein yes scroll yes rows 4 required file testC*.shp -FILE 'Please select a set of test data '
%2 input .licenseshp 32 typein yes scroll yes rows 4 required file testl*.shp -FILE 'Please select a license data'
%3 input .river 32 typein yes scroll yes rows 4 required file testr*.shp -FILE 'Please select a river data'
%4 input outcov 32 init # help 'Default is <in_cover>' character
%5 input .idealarea 8 size 8 init 1000000 choice 200000 500000 1000000 1500000 -prompt 'Select the desired district size'
%6 input .size 4 size 4 init 500 choice 100 250 500 1000 2000 5000 -prompt 'Select the desired triangular size'
%7 input .weight1 4 size 4 init 7 choice 1 3 5 7 9  -prompt 'Select fuzzy weight for Euclidean Measure'
%8 input .weight2 4 size 4 init 1 choice 1 3 5 7 9  -prompt 'Select fuzzy weight for Fractal Dimension'
%11 input .weight3 4 size 4 init 7 choice 1 3 5 7 9  -prompt 'Select fuzzy weight for Application Dependent Factor'
%9 input .confidence 4 size 4 init 0.5 choice 0 0.1 0.3 0.5 0.7 0.9 1.0  -prompt 'Select your degree of confidence'
%10 input .risk 4 size 4 init 0.5 choice 0 0.5 1  -prompt 'Select optimistic, moderate or pessimistic view'
%ok button OK &return; &r preprocess1.aml %outcov%
%help button return keep 'HELP' help clean
%cancel button cancel 'CANCEL' &return
```

```
*************************************************************************************************
* FILE NAME: Preprocess1.aml    PURPOSE: To prepare data for redistricting process, partial process for the DP module
*************************************************************************************************
&arg name
&term 9999
display 9999

&thread &create 'Status' &menu message ~
  &position &cc &screen &cc
&sv .message = Preparing Fuzzy Weight Value
&THREAD &SYNCHRONIZE  status
&r pre_weight
&sv .message = Start processing...
&THREAD &SYNCHRONIZE  status
&if [exists splitcov -cover] &then;  kill splitcov
&if [exists tintemp -tin] &then;  kill tintemp
&if [exists covtemp -cover] &then;  kill covtemp
&if [exists shparc -cover] &then;  kill shparc
&if [exists license -cover] &then;  kill license
/*---------------- Preprocessing Stage:  Get input source, license source and  triangles size  from user.
/* Then convert source to relevant coverage and create the triangles by CREATETIN  ---------------
shapearc %.inshp% shparc
shapearc %.licenseshp% license
clean license
build license poly
&sv .message = Pre-processing now...
&THREAD &SYNCHRONIZE  status
createtin tintemp %.size%
cover shparc line elev_meter
end
describe tintemp
tinarc tintemp covtemp poly degree
intersect covtemp license splitcov poly
dropitem splitcov.pat splitcov.pat covtemp# covtemp-id license# license-id
/*---------------find the stage interval----------------------------------------------
&if [exists begin -info] &then;   killinfo begin
ae
mape splitcov
edit splitcov poly
select all
unselect degree_slope = -9999
statistic # begin
max degree_slope; min degree_slope; mean degree_slope; std degree_slope
~
n
n
&sv stdd = [listunique begin -info STD-DEGREE_SLOPE]
&sv stdd = [ round %stdd% ]
&sv stddd = %stdd% / 2
&sv stddd = [ round %stddd% ]
q
&sv .message = Start process...
&THREAD &SYNCHRONIZE  status
&s split := splitcov
dropitem %split%.pat %split%.pat SAREA
additem %split%.pat %split%.pat V 8 18 F 4
&if [exists temp -cover] &then;   kill temp
&if [exists w -cover] &then;   kill w; &if [exists x -cover] &then;   kill x
&if [exists y -cover] &then;   kill y; &if [exists z -cover] &then;   kill z
/*----Divide Stage: Get triangles with slope greater than 35 degree and update the result with  the original coverage
&sv .message = Processing slope > %stdd%
&THREAD &SYNCHRONIZE  status
reselect  %split% z
```

```
reselect degree_slope gt %stdd%
~
n
n
&sv cover = z
&r process.aml %cover%; &r pre_value %cover%
reselect z z1
reselect V >= 0.3 and degree_slope ge 0 /*and area >= %allowminarea% and area <= %allowmaxarea%
~
n
n
/*------------------------ Check result of worst case------------------------
&if [exists z1.tab -info] &then; killinfo z1.tab
statistics z1.pat z1.tab
sum area
~
n
n
&sv adaz = [ listunique z1.tab -info FREQUENCY]
&if %adaz% = 1 &then
    &do
     &sv .message = Worst Case do not produce compact district;      &THREAD &SYNCHRONIZE status
    &end
&else
    &do
     update splitcov z1 temp poly # keepborder
     kill splitcov
     copy temp splitcov
    &end
/*--Divide Stage: Get triangles with slope < 35 AND slope > 25 degree and update the result with original coverage
&sv .message = Processing slope > %stddd% and slope < %stdd%
&THREAD &SYNCHRONIZE status
reselect %split% y
reselect ( degree_slope > %stddd% ) or v >= 0.3
~
n
n
&sv cover = y
&r process.aml %cover%
&r pre_value %cover%
/*kill temp
reselect y y1
reselect V >= 0.3 and degree_slope ge 0 /* V >= 0.3 and area >= %allowminarea% and area <= allowmaxarea%
~
n
n
/*------------------------ Check result of worst case------------------------
&if [exists y1.tab -info] &then; killinfo y1.tab
statistics y1.pat y1.tab
sum area
~
n
n
&sv aday = [ listunique y1.tab -info FREQUENCY]
&if %aday% = 1 &then
    &do
     &sv .message = Second Worst Case do not produce compact district
     &THREAD &SYNCHRONIZE status
   &end
&else
    &do
     &if [exists temp -cover] &then; kill temp
     update splitcov y1 temp poly # keepborder
```

```
    kill splitcov
    copy temp splitcov
  &end
/*-----------Divide Stage: Get triangles with slope < 25 and Fuzzy value < 0.5 to create combination district late------
&sv .message = Processing slope <= %stddd%
&THREAD &SYNCHRONIZE status
reselect %split% x
reselect  V <= 0.3   /* v <= 0.3
~
n
n
&sv cover = x
&r process.aml %cover%
&call Allow
&return
/*************************
&routine Allow  /*---Prompt user to get permission for combination district ----------
/*************************
&sv .message = Start conquering...
&THREAD &SYNCHRONIZE status
&if [exists temp -cover] &then;  kill temp
&if [exists %name% -cover] &then;  kill %name%
&sv .message = Final conquer...
&THREAD &SYNCHRONIZE status
&if %aday% <> 1 and %adaz% <> 1 &then
 &do
   update x y1 temp poly # keepborder
   update temp z1 %name% poly # keepborder
 &end
&if %aday% = 1 and %adaz% = 1 &then;   copy x %name%
&if %aday% <> 1 and %adaz% = 1 &then;   update x y1 %name% poly # keepborder
&r postprocess1 %name%
   arcplot
   mape %name%
   polygons %name%
&sv .message = Process Finish
&THREAD &SYNCHRONIZE status
&return

*************************************************************************************************
* FILE NAME    : Process.aml      PURPOSE        : Start process for each slope range
*************************************************************************************************
&arg cover
&severity &error &routine error
INDEX %cover% POLY CREATE
&sv allow = 0
&sv eacharea = %.idealarea%
ae
 mape %cover%
 edit %cover% poly
 de polygon
 draw
 &sv end = [show number total]
 &call Start
 draw
 save
q
tables
select %cover%.pat
&select %cover%
 &when z ;   calculate aspect = 3
 &when y ;  calculate aspect = 2
 &otherwise ;    calculate aspect = 1
```

```
&end
quit
&return

/*************************
&routine Start          /* --------------Start Redistrict-----------------------------------------------
/*************************
&do n = 1 &repeat %n% + 1 &until %n% = %end% /* for loop
    clear
    /*-------- unselect degree_slope = -9999 so that time will not waste to do for other range ---
    select for $Recno = %n% AND degree_slope <> -9999
    &sv count = [show number select]
    &if %count% <> 0 &then  /*  if 1
    &do
      selectput arcplot
      arcplot
      &sv area = [listunique %cover% -poly area]
      &if %area% < [abs [calc %eacharea% ]] &then /* if 2
      &do
         aselect %cover% POLY ADJACENT 0
         /*---------- fuzzy value for the idealarea -----------------------
         &sv idealarea = [abs [calc %eacharea% ]]
         quit
         selectget
         /*unselect degree_slope = -9999 so that adjacent will not involve process duplication
          unselect  degree_slope = -9999
         &sv adjacentcount = [show number select]
         &if %count% <> %adjacentcount% &then
           &call Merge
     &end
     &else
       q
   &end /* end if 1
&end /* end for loop
  select area < %eacharea%
  &if %end% <> [show number select] AND  [show number select] <> 0 &then
  &do
    &sv end = [show number select]
    &call Start
  &end
&return
/*************************
&routine Merge
/*************************
   drawselect
   &call Drop
   &if %allow% = 2 &then
   &do
     &sv allow = 0
     &return
   &end
    merge select
   &if [show number select] = 0 &then
   &return
    calculate degree_slope = %slope%
    save
    &label AGAIN
&return

/*************************
&routine Drop
/*************************
&if [exists statarea -info] &then;  arcplot arc killinfo statarea
```

```
&if [exists statsumarea -info] &then;  arcplot arc killinfo statsumarea
/*------Find sum of adjacent polygons --------------------------------
statistics
N
sum area
max degree_slope
END
statistics # statsumarea
Y
&sv slope = [listunique statsumarea -info MAX-degree_slope]
&sv sum_area = [listunique statsumarea -info SUM-area]
&sv remain =  [show number select]
&if %sum_area% <= %eacharea%  &then
     &return
&else
&do
   &if %remain% <= 2 &then
    &do
      /* fuzzy value for the idealarea the relative measure
       &sv relative = [abs [calc %area% / %sum_area% ]]
       &if %relative% < 0.20 &then
         &do
              &return
         &end
       &else
         &do
              &sv allow = 2
           &return
        &end
    &end
    unselect for $Recno =  %n%
    statistics
    N
    max area
    min area
    END
    statistics # statarea
    Y
    &sv max_area = [listunique statarea -info MAX-AREA]
    &sv min_area = [listunique statarea -info MIN-area]
   /*---- Drop Polygon if the sum area of the adjacent exceed the ideal area  -------------------
     &if %max_area% >= [abs [calc %eacharea% / 3 ]] &then
       unselect area gt %max_area%
     &else
        reselect area gt %min_area%
        aselect for $Recno =  %n%
     &call Drop
&end
&return
/*********************
&routine error
/*********************
 &goto AGAIN; &return  Cannot merge due to adjacent with common node
```

```
*****************************************************************************************************
* FILE NAME          : Pre_weight.aml       PURPOSE                : prepare the weighting vector
*****************************************************************************************************
&sv .w1 := %.weight1%
&sv .w2 := %.weight2%
&sv .w3 := %.weight3%
&if %.w1% = 1 &then;     &do;  &sv .w11 := 1; &sv .w12 := 1; &sv .w13 := 3; &end
&if %.w1% = 3 &then;     &do;  &sv .w11 := 1; &sv .w12 := 3; &sv .w13 := 5; &end
&if %.w1% = 5 &then;     &do;  &sv .w11 := 3; &sv .w12 := 5; &sv .w13 := 7; &end
```

```
&if %.w1% = 7 &then;     &do;  &sv .w11 := 5; &sv .w12 := 7; &sv .w13 := 9; &end
&if %.w1% = 9 &then;     &do;  &sv .w11 := 7; &sv .w12 := 9; &sv .w13 := 9; &end
&if %.w2% = 1 &then;    &do;  &sv .w21 := 1; &sv .w22 := 1; &sv .w23 := 3; &end
&if %.w2% = 3 &then;    &do;  &sv .w21 := 1; &sv .w22 := 3; &sv .w23 := 5; &end
&if %.w2% = 5 &then;    &do;  &sv .w21 := 3; &sv .w22 := 5; &sv .w23 := 7; &end
&if %.w2% = 7 &then;    &do;  &sv .w21 := 5; &sv .w22 := 7; &sv .w23 := 9; &end
&if %.w2% = 9 &then;    &do;  &sv .w21 := 7; &sv .w22 := 9; &sv .w23 := 9; &end
&if %.w3% = 1 &then;    &do;  &sv .w31 := 1; &sv .w32 := 1; &sv .w33 := 3; &end
&if %.w3% = 3 &then;    &do;  &sv .w31 := 1; &sv .w32 := 3; &sv .w33 := 5; &end
&if %.w3% = 5 &then;    &do;  &sv .w31 := 3; &sv .w32 := 5; &sv .w33 := 7; &end
&if %.w3% = 7 &then;    &do;  &sv .w31 := 5; &sv .w32 := 7; &sv .w33 := 9; &end
&if %.w3% = 9 &then;    &do;  &sv .w31 := 7; &sv .w32 := 9; &sv .w33 := 9; &end
  &ty %.w11% %.w12% %.w13% ;   &ty %.w21% %.w22% %.w23%;   &ty %.w31% %.w32% %.w33%
&return

***************************************************************************************************
*FILE NAME          : Pre_value..aml        PURPOSE  : To prepare the Fractal Dimension and Euclidean Measure
***************************************************************************************************
&args complexcov
INDEX %complexcov% POLY CREATE
dropitem %complexcov%.pat %complexcov%.pat FD EM
&r killmany bdy*; &r killmany line*; &r killmany poly*
&if [exists totalV -info] &then; killinfo totalV
/*--------------Create the grid (for polygon and line of the district ) to find Fractal Dimension----------------
build %complexcov% line
grid
&do i  = 4 &repeat %i% + 4 &until %i% = 20
  poly%i% = polygrid(%complexcov%, %complexcov%#, #, #, %i%)
  line%i% = linegrid(%complexcov%, Rpoly#, #, #, %i%)
  bdy%i% = selectmask( poly%i%, line%i%)
&end
quit
/*------------------ Prepare log-log value to plot graph to find the slope (slope = Fractal Dimension)------------------
&do i  = 4 &repeat %i% + 4 &until %i% = 20
 &r calc_log BDY%i%.VAT COUNT NEWITEM log%i%
 TABLES
 SELECT BDY%i%.VAT
 REDEFINE
 1
 %complexcov%#
 4
 10
 B
 ~
 Q
 /*-----------drop unnecessary items and join the relevant features file ---------------------
 dropitem bdy%i%.vat bdy%i%.vat count
 joinitem %complexcov%.pat bdy%i%.vat %complexcov%.pat %complexcov%# # ORDERED
 dropitem %complexcov%.pat %complexcov%.pat value
&end
/*--------------------- Prepare value to get slope ---------------------------------------------------------------------------------
&sv X = [CALC [LOG10 20] + [LOG10 4] + [LOG10 8] + [LOG10 12] + [LOG10 16] ]
&sv X2 = [CALC [LOG10 20] * [LOG10 20] + [LOG10 4] * [LOG10 4] + [LOG10 8] * [LOG10 8] + [LOG10 12] *
[LOG10 12] + [LOG10 16] * [LOG10 16] ]
tables
/*----Calculate Eucliden measures value (EM) and Fractal dimension value (FD). Then insert to the features table---
additem %complexcov%.PAT EM 8 8 f 4
ADDITEM %complexcov%.PAT FD 8 8 F 4
ADDITEM %complexcov%.PAT TEMP 8 8 F 4
select %complexcov%.PAT
calculate EM = ( Perimeter * Perimeter ) / Area
CALCULATE TEMP = ( ( LOG20 * [LOG10 20] ) + ( LOG4 * [LOG10 4] ) + ( LOG8 * [LOG10 8] ) + ( LOG12 * [LOG10
12] ) + ( LOG16 * [LOG10 16] ) )
```

```
CALCULATE FD = ( LOG20 + LOG4 + LOG8 + LOG12 + LOG16 )
CALCULATE FD = ( ( 5 * TEMP ) - ( FD * %X% ) )
CALCULATE FD = ( 0 - ( FD / ( [calc 5 * %X2% ] - ( [calc %X% * %X% ] + 0.00000001 ) ) ) )
DROPITEM %complexcov%.PAT TEMP log20 log4 log8 log12 log16
/*-------------- Prepare Fuzzy weight value. ------------------------
&r pre_Fuzzy %complexcov%
q
dropitem %complexcov%.pat %complexcov%.pat EM FD
&return END

*****************************************************************************************************
* FILE NAME          : Pre_fuzzy..aml      PURPOSE    : To produce Integrated Compactness Index (ECI)
*****************************************************************************************************
&args complexcov
&severity &error &routine error
&term 9999
additem %complexcov%.pat c11 4 4 f 2;additem %complexcov%.pat c12 4 4 f 2
additem %complexcov%.pat c13 4 4 f 2;additem %complexcov%.pat c21 4 4 f 2
additem %complexcov%.pat c22 4 4 f 2;additem %complexcov%.pat c23 4 4 f 2
additem %complexcov%.pat c31 4 4 f 2;additem %complexcov%.pat c32 4 4 f 2
additem %complexcov%.pat c33 4 4 f 2;additem %complexcov%.PAT c1 8 8 f 4
additem %complexcov%.PAT c2 8 8 f 4;additem %complexcov%.PAT c3 8 8 f 4
additem %complexcov%.PAT A1 8 8 f 4;additem %complexcov%.PAT A2 8 8 f 4
additem %complexcov%.PAT V0 8 8 f 4;additem %complexcov%.PAT V2 8 18 f 4
select %complexcov%.PAT
reselect em > 30; calculate c11 = 1;calculate c12 = 1;calculate c13 = 3; select %complexcov%.PAT
reselect em > 25 and em <= 30; calculate c11 = 1;calculate c12 = 3;calculate c13 = 5; select %complexcov%.PAT
reselect em >= 20 and em < 25;calculate c11 = 3;calculate c12 = 5;calculate c13 = 7; select %complexcov%.PAT
reselect em >= 15 and em < 20;calculate c11 = 5;calculate c12 = 7;calculate c13 = 9; select %complexcov%.PAT
reselect em < 15;calculate c11 = 7;calculate c12 = 9;calculate c13 = 9; select %complexcov%.PAT
reselect fd >= 1.01 or fd < 0.8;calculate c21 = 1;calculate c22 = 1;calculate c23 = 3; select %complexcov%.PAT
reselect fd < 1.01 and fd >= 1.08;calculate c21 = 1;calculate c22 = 3;calculate c23 = 5; select %complexcov%.PAT
reselect fd < 1.08 and fd >= 1.05;calculate c21 = 3;calculate c22 = 5;calculate c23 = 7; select %complexcov%.PAT
reselect fd < 1.05 and fd >= 1.00;calculate c21 = 5;calculate c22 = 7;calculate c23 = 9; select %complexcov%.PAT
reselect fd < 1.00 and fd >= 0.8;calculate c21 = 7;calculate c22 = 9;calculate c23 = 9; select %complexcov%.PAT
reselect ( area < [ calc %.idealarea% * 0.7 ] ) or ( area > [ calc %.idealarea% * 1.3 ] );
calculate c31 = 1;calculate c32 = 1;calculate c33 = 3
select %complexcov%.PAT
reselect ( area >= [ calc %.idealarea% * 0.7 ] AND area < [ calc %.idealarea% * 0.8 ] ) or ( area <= [ calc %.idealarea% * 1.3
] AND area >  [ calc %.idealarea% * 1.2 ] )
calculate c31 = 1;calculate c32 = 3;calculate c33 = 5
select %complexcov%.PAT
reselect ( area >= [ calc %.idealarea% * 0.8 ] AND area < [ calc %.idealarea% * 0.9 ] ) or ( area <= [ calc %.idealarea% * 1.2
] AND area >  [ calc %.idealarea% * 1.1 ] )
calculate c31 = 3;calculate c32 = 5;calculate c33 = 7
select %complexcov%.PAT
reselect ( area >= [ calc %.idealarea% * 0.9 ] AND area < %.idealarea% ) or ( area <= [ calc %.idealarea% * 1.1 ] AND area
> %.idealarea% )
calculate c31 = 5;calculate c32 = 7;calculate c33 = 9
select %complexcov%.PAT
reselect area = %.idealarea%
calculate c31 = 7;calculate c32 = 9;calculate c33 = 9
select %complexcov%.PAT
calculate c1 = c11 * %.w11% + c21 * %.w21%  + c31 * %.w31%
calculate c2 = c12 * %.w12% + c22 * %.w22%  + c32 * %.w32%
calculate c3 = c13 * %.w13% + c23 * %.w23%  + c33 * %.w33%
calculate A1 = ( c2 - c1 ) * %.confidence% + c1
calculate A2 = ( c2 - c3 ) * %.confidence% + c3
calculate V0 = %.risk% * A2 + ( 1 - %.risk% ) * A1
calculate V2 = V0 * V0
statistic # totalV
sum V2
end
```

```
&sv grandtotal = [listunique totalV -info SUM-V2]
&sv grandtotal = [sqrt %grandtotal%]
&ty %grandtotal%
calculate V = V0 / %grandtotal%
dropitem %complexcov%.pat C11 C12 C13 C21 C22 C23 C31 C32 C33
dropitem %complexcov%.pat c1 c2 c3
dropitem %complexcov%.pat A1 A2 V0 V2
&return END
/********************
&routine error
/********************
&return Error

**************************************************************************
* FILE NAME         : Postprocess1.aml     PURPOSE : To consider the natural boundary
**************************************************************************
&arg cover
&if [exists %cover%P -cover] &then; kill %cover%P
&if [exists elicov2 -cover] &then; kill elicov2
&if [exists intcov -cover] &then; kill intcov
&if [exists unioncov -cover] &then; kill unioncov
&if [exists river -cover] &then; kill river
shapearc %.river% river
intersect river %cover% intcov line
build intcov poly
union intcov %cover% unioncov # join
dropitem unioncov.pat unioncov.pat intcov-id intcov# %cover%# %cover%-id
build unioncov line
build unioncov poly
eliminate unioncov %cover%P keepedge line
reselect length > 0
~
n
n
joinitem %cover%P.aat intcov.aat %cover%P.aat length
build %cover%P line
build %cover%P poly
ae
mape %cover%P
edit %cover%P arc
de arcs
draw
select rpoly# = 1 or lpoly# = 1
drawselect
aselect type = 'R'
drawselect
calculate %cover%P-id = -1
nselect
drawselect
generalize 300 pointremove
drawselect
select type = 'R'
calculate %cover%P-id = 1
Save
&if [exists statsumarea -info] &then; arcplot arc killinfo statsumarea
mape license
edit license poly
select all
statistics
N
sum area
END
statistics # statsumarea
```

```
Y
&sv sumarea = [ listunique statsumarea -info SUM-AREA]
q
clean %cover%P
joinitem %cover%P.pat unioncov.pat %cover%P.pat  %cover%P# # link
dropitem %cover%P.pat %cover%P.pat  unioncov# unioncov-ID
kill %cover%
&r pre_value  %cover%P
&if [exists p.tab -info] &then;   killinfo p.tab
statistics %cover%p.pat p.tab
sum area
mean V
~
n
n
&sv adap = [ listunique p.tab -info FREQUENCY]
&sv MeanV = [ listunique p.tab -info MEAN-V]
&ty %adap% %sumarea%  [round [calc %sumarea% / %.idealarea%]]
&if %adap% >= [round [calc %sumarea% / %.idealarea%]] &then
&do
  eliminate %cover%P  %cover%
  reselect  V < [ calc %MeanV% * 0.8 ] or area < [calc %.idealarea% / 3 ]
  ~
  n
  n
 &end
&r pre_value  %cover%
&return
```

```
**********************************************************************************************
* File name            : killmany.aml         PURPOSE: Kills multiple Arc/Info data sets at the same time.
**********************************************************************************************
&args .template .option
/*---------------------check arguments--------------------------------------
&if [null %.option%] &then &set .option = arc
&if [null %.template%] &then
  &do
    &type purpose: kills multiple geographic data sets
    &type         (coverages or grids) which match a   user-specified template.
    &type usage: &r killmany <template> {arc | info | all}
    &type       where <template> = dataset root name and *
    &return
  &end
&set covers [listfile %.template% -cover]
&if [null %covers%] &then
  &do
    &type there are no coverages that match the specified template
    &goto grids
  &end
kill (!%covers%!) %.option%
&label grids
&set grids [listfile %.template% -grid]
&if [null %grids%] &then
  &do
    &type there are no grids that match the specified template
     &return
  &end
kill (!%grids%!) %.option%
&return
```

APPENDICES B:
RELATED CONFERENCE PAPER

The 3rd ISPRS workshop on Dynamic and Multi-dimensional GIS, The 10th Annual Conference of CPGIS on Geoinformatics, From May 23-25, 2001, at Asian Institute of Technology, Bangkok, Thailand

Integration Of Compactness Measurement Methods Using Fuzzy Multicriteria Decision Making: A New Approach For Compactness Measurement In Shape Based Redistricting Algorithm

Yin Chai WANG
Faculty of Information Technology, Universiti Malaysia Sarawak,
94300 Kota Samarahan, Sarawak, MALAYSIA
Tel:+60-82-671000 ex377, Fax:+60-82-672301, E-mail:ycwang@fit.unimas.my

Chin Wei BONG
Faculty of Information Technology, Universiti Malaysia Sarawak,
94300 Kota Samarahan, Sarawak, MALAYSIA

ABSTRACT

Thousand of district plans are generated from the redistricting process but most of the processes are unable to determine the best alternative. This paper will discuss the limitation of existing redistricting algorithms especially on redistricting applications based on the geographical compactness which seems to have great limitation because it causes the redistricting process more complex and intractable. Spatial complexity for redistricting like topography factors causes inflexibility on conventional computer and also difficulties was found especially in spatial aspect like geographical size, district shape and boundary lines which bring to the issues of compactness and continuity. This paper presented a new approach for compactness measurement in a shape based redistricting planning by using fuzzy multiple criteria decision-making to enhance the redistricting process especially on district compactness and continuity. The design and development process of the proposed approach will also highlighted in this paper including generation of compactness index based on the synthesis of the concepts of fuzzy set theory, AHP, α-cuts concept and index of optimism of district planners to estimate the degrees of satisfaction of the judgements on a district plan. The performances of the proposed new approach was tested on a forest blocking prototype to demonstrate its applicability in redistricting applications with respect to their redistricting goals and criteria. The results shown the proposed method is more flexible, simple and comprehensive with easy computation and efficiency which facilitates its uses in compactness measurement in redistricting applications.

B.1.1. INTRODUCTION

According to Altman [2], Ronald Reagan was not the only recent academic to state that computer could remove the controversy from redistricting and it can find the "optimal" redistricting plan by given any set of values that can be specified. In order to draw a fair and unbiased district, the methods and algorithms concerned are playing an important roles and none of the methods is specifically being determined as a standard. Although redistricting is an optimization problem, different solutions on integer programming or linear programming are still did not offer the best solution. The geographical features play an important role and the measures of geographical shape will be an enormous endeavor in enhancing the redistricting result from the application dependent factor. However, traditional redistricting measures on spatial features are unsatisfactory to produce the optimal result in terms of their compactness and continuity. Thus, few effort to create an optimum district boundary has been proposed and one of the idea is to use fractal dimension to measure the spatial features in the redistricting application.

Previous efforts on redistricting can be analyzed in two perspectives of a district plan. Hence, district plan is defined as the output of the redistricting process. Firstly, it is about the standards to evaluate the district plan. These standards refer to the compactness measurement techniques. Second, it is on the redistricting methods, techniques or algorithms in creating district plan. The existing standard to evaluate and assess the district plan is use compactness measurements. These compactness measurements encompass continuity and therefore it means that the maximally compact plan will not measurably and avoidably noncontiguous [2]. Therefore, although the evaluation tools are called compactness measurements, they assess not only on the compactness but also on the continuity of the districts in the plan. Next section will discuss in details the problems and limitation of compactness measurement methods for shape based redistricting system.

The primary concern on the proposed redistricting system is on the shape regularity and the continuity of the district. The main component in the proposed method is to use GIS tool in preprocessing the necessary input data, assisting the redistricting process and producing an optimum result. The main component should be to improve the compactness measurement index through enhancement of the decision making process and dynamic programming. On the other hand, the method used to redraw the redistricting boundary in this research is the FMCDM methods as to draw optimal district shape or to get optimal compact district plan. Therefore, the compactness measurement technique will be used as criteria in measuring the shape compactness or regularity and continuity. Although there are many compactness measurement methods, MCDM may enable more than two methods being used in the integrated model. Besides, the advanced OR technique with the fuzzy set theory is believed to be useful and helpful in vague compactness measurement index. For implementation purposes, Dynamic programming will be promising for the optimality in the enhanced model

B.1.2. COMPACTNESS MEASUREMENT LIMITATION & PROBLEMS

The various types of compactness measurement for evaluating the result of the redistricting plan are from tremendous endeavor of compactness measurements in redistricting. It really shows the importance of geographical aspect in redistricting and how necessary for different techniques are to be used to measure the compactness and continuity of the district plans. However, these compactness measurements commonly used after the redistricting process and usually treated as separately to redistricting effort. The reason for assessing compactness and continuity of the redistricting plan is to ensure balance and fair result. However, current research does not help district planners to determine which compactness measures to use, which are effective, and neutral. Therefore, choosing a particular compactness measure was a special challenge because previous researchers have proposed over thirty distinct measures of compactness.

Most of the existing redistricting methods do not primarily concern on the shape or spatial context and suffering from problems in implementation because lack of systematic evaluation, inferior decision making or decision support process. Compactness measurement for district compactness and continuity assessment gives an index for assessment and for decision making of compact district but these methods do not permit at imprecision and partial occurrence, which is not compatible with actual decision making process. Frequently, decision making in spatial assessment involves uncertainty due to incompleteness of the data acquired and the variety of evaluation tools used to gather data. In addition, numerical measurements in compactness measurements give uncertain and vague value for accurate assessment of district compactness and continuity. Altman's axiomatic assessment on most of the compactness measurements show that existing compactness measurement does not accurately assess compact district shape.

There is a lot of redistricting techniques but still does not give an ideal district plan in term of geographical aspect. Redistricting cannot run away from compactness and continuity. Many redistricting problems suffered for the spatial context problem. However, there is no incorporation of compactness measurement to redistricting techniques to get optimal compact district. Shape-based redistricting is a multiple criteria problem because redistricting is done for application dependent criteria but at the same time, application independent criteria shouldn't be ignored. Thus, multiple criteria are to be considered to draw a district plan. General redistricting, which aims to maintain optimal compactness, continuity and other application dependent criteria is a complex multiple criteria problem.

The complexity of the data structures and the data volumes are the also common problems to be solved in redistricting arena. Due to the complexity of the data structure, there is an insufficiency in collecting relevant data. For instance, in the political redistricting, there is no single source contains all geographical and population data for the United States congressional district over the entire period from 1789 through 1912 [1]. The complexity of data structures needs certain kind of solution that is able to assign and cope with different kind of behavior like the spatial and non-spatial behavior, consider their relationships and dependencies in order to ease the manipulation of each part. For instances, in order to consider the non-spatial behavior in redistricting, there is a necessity to find solution where it is possible to store, manipulate, visualize, and relate data especially for other spatial behavior. It is certainly that redistricting will not be done by only single type of data. Therefore, the relationship between data and behavior plays an important role.

In order to get optimum shape in redistricting; compactness measurement will be necessary to measure the regularity of the district shape. Therefore, the research will propose to use these compactness measurements as goals to draw district instead of using them to assess the compactness after redistricting process.

Consequently, this subsection aims to mention and summarize the shape based redistricting problem in order to enhance existing redistricting model to get optimal compact district plan by proposed integrated model for shape based redistricting as discussed in the next section.

B.1.3. COMPACTNESS MEASUREMENT INDEXING USING FUZZY MULTICRITERIA DECISION MAKING

B.1.3.1 The Design and Model

The proposed system aims to get optimum or compact district, which defined as in this research is as regular and contiguous as possible and at the same time to use the natural boundaries where possible in the district line. The main significance is to ease the control of a space. For example, with regular and continuous districts, governors face less difficulty in ruling a district or even country in political redistricting.

In order to solve the problems mentioned, the system design must be able to fulfill the requirements as stated below:

- To produce better or enhanced shape assessment index which is more descriptive and able to incorporate with natural feelings of district planners or decision makers. Natural feelings here may include their confidence and their attitude to risk.
- To cope with fuzziness in the shape assessment index.
- To integrate multiple compactness measurement method in order to gather the strengths of particular method and at the same time to reduce or minimize its weaknesses or lacks.
- To incorporate the new index into redistricting process or algorithm to generate an optimal compact district
- To work and perform in an environment (software or hardware interface) that is able to manage the spatial and non-spatial data, their relationship and dependency

The domain and task of the proposed model is at enhancing the decision support system in maintaining the compactness and continuity of the shape based redistricting to produce the optimal compact district plan (the output). Fuzzy Multi Criteria Decision Making (FMCDM) will be used to enhance this component of the model. GIS analysis tools are used as important supportive tools for this model because of its capability to due with relationships of spatial and non-spatial data. Besides, it also provides numerous ready-to-use functions for a shape based redistricting.

The following subsections describe the proposed design model for the shape based redistricting by integrating based on the requirement mentioned as shown in Figure B.1.1. The components of the proposed integrated model is Application Independent Data Store (AID), Application Dependent Data Store (ADD), Shape Optimal Rules (SOR), Data Preparation Module (DPM), Fuzzy Multi Criteria Decisions Making (FMCDM), Combine Optimal District Module (COD) and stages of decision making.

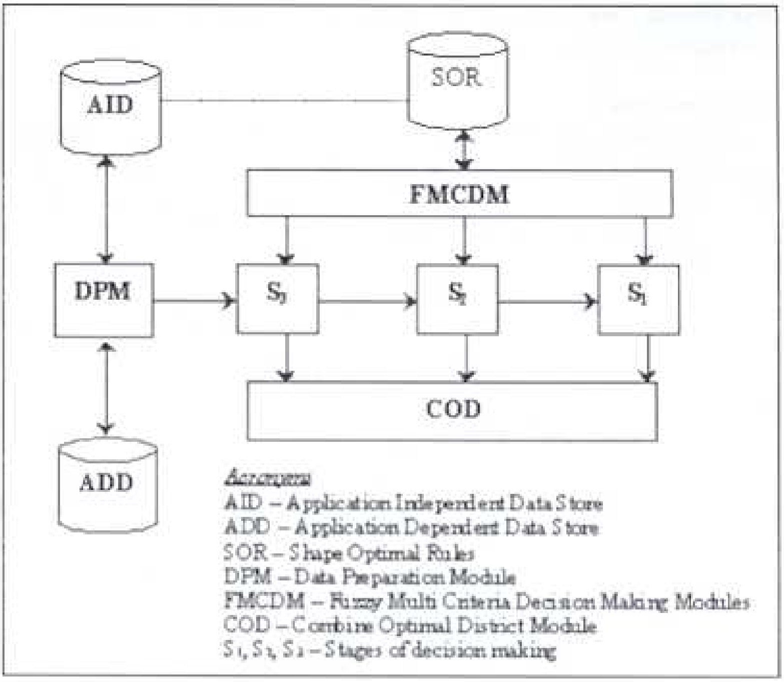


Figure B.1.1: Integrated Compactness Measurement Using Fuzzy Multiple Criteria Decision Making

Data Preparation Module (DPM) will prepare data in triangulated irregular network format to ensure the model can work with appropriate datasets. A triangulated irregular network is made up of arcs, which define the boundary, and a label point, which links the polygon feature to an attribute record. As the mother of all polygon partitioning problems is triangulation, the interior of all kind of polygons can be completely partitioned into triangles [6].

B.1.3.2 Applying Fuzzy Set Theory for Compactness Measurement in the SOR

Shape Optimal Rules play an important role for the standard of the optimality of district shape. This component is related to the application independent data because the focus of the research is on the geographical aspect.

In this research study, two compactness measurement methods are selected to gather the district boundary complexity and district compactness, which is Euclidean measure and non-Euclidean measure. More than one measurements are used in this research aims to combine the strengths of each measurement to the district compactness and to reduce their weaknesses by the weighing vector which will be incorporated into the enhanced index. The compactness measurement index for both of the methods give a gray area (undecided) for district planners or decision makers on the district shape assessment. Therefore, fuzzy set theory as will be used to formalize the gray area in order to specifically express and also never ignore the gray area.

Linguistic terms which have been found intuitively easy to use [4] are used to represent the decision maker's subjective assessment on the compactness measurement index. A term set of linguistic for each compactness measurement index is {Very Poor (VP), Poor (P), Fair (F), Good (G), Very Good (G)}, defined in Figure B.1.2. The weighing vector for the compactness evaluation criteria will be directly given by the decision maker or obtained by using pairwise comparison. Meanwhile, the term set used for the weighing vector is {Least Important (LTI), Less Important (LSI), Important (I), More Important (MEI), Most Important (MTI), Most Important (MTI)}, defined as in Figure B.1.3.

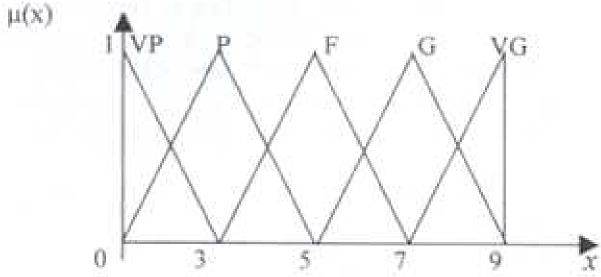


Figure B.1.2.: Membership functions, μ(x) and the linguistics terms for CDI, x

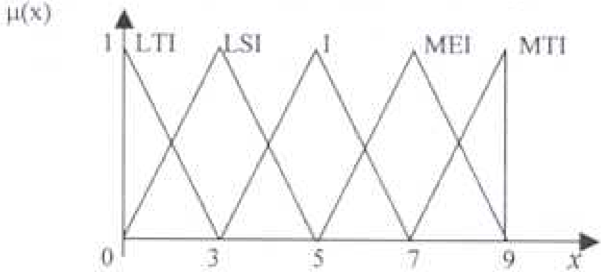


Figure B.1.3 : Membership function, μ (x) and the linguistic terms used by the weighing vector, x

B.1.3.3 Decision Support with FMCDM

The integrated model will able to consider multiple compactness measurements as a measurement instruments to ensure optimal compact district. However, the compactness measurement provides numerical index, which is vague and may be incomplete. Therefore, Fuzzy Multiple Criteria Decision Making (FMCDM) will be able to provide a solution for the problem mentioned. Therefore, multiple compactness measurement will used to produce an integrated compactness index that is able to cope with fuzziness to measure the geographical aspect of the district plan. Multi Criteria Decision Making (MCDM) deals with problem of helping the decision maker to choose the best alternatives, according to several criteria (Valls, 2000). This approach will be used to solve complex shape based redistricting problems in a systematic, consistent and more productive way because it may enhance the degree of conformity and coherence in the decision process [3]. Meanwhile, the fuzziness concept is used to solve the subjectiveness and vagueness for deciding the compactness assessment index. Therefore, fuzzy sets theory to MCDM models will be used to provide an effective way of dealing with the subjectiveness and vagueness of decision making process for the general multiple criteria shape based redistricting problems. This method is to support a systematic decision making for several reasons [7]. First, the information and knowledge for the redistricting decisions is incomplete, uncertain or imprecise or even inconsistent state clearly for this but the information overload is still increasing. Second, there are also multiple conflicting goals and multiple different type of constraint.

Therefore, the proposed redistricting environment will be the integration of FMCDM in GIS environment to enhance the shape-based redistricting process. Indeed, FMCDM is an Operation Research (OR) technology [5] that can face the complexity of the environment, which strategic decisions are needed especially like the redistricting problems. In multiple criteria program, redistricting application functions are established to measure the degree of fulfillment of the decision maker's requirements about the goal function and are extensively used in the process of finding "good compromise" solution [3]. For district planners, the requirements in redistricting application include the achievement of goals on compactness, nearness to an ideal point on the application dependent factor, and other satisfaction. According to Fuller and Carlsson, one of the earliest practical applications of FMCDM is a commercial application for evaluation of the credit-worthiness of credit card applicants [3]. Besides, the recent applications including the evaluation of weapons systems, a project maturity evaluation system at Mercedes-Benz in Germany, technology transfer strategy selection in biotechnology and aggregation of market research data.

The FMCDM method to be used in this research is Fuzzy Analytical Hierarchy Process or Fuzzy AHP. Analytical Hierarchy Process is a multi criteria method, which uses hierarchic structures to represent a decision problem, and then develops priorities for the factors based on decision maker's judgment. It has been widely used to solve complicated, unstructured decision problem and thus it should be concerned with the processing of fuzzy information. As it is difficult to get exact ratios for a pair of factors considered, a fuzzy ratio for the relative significant may incorporates the natural feelings of human beings. In other words, fuzzy theory is effective when the situation contains fuzziness from human subjectivity in redistricting functions. Indeed, Fuzzy AHP method was discussed and being used for ranking of Indian Coals in industrial use [8], rate and ranking the disability [9] and also assessing risk of cumulative trauma disorders [10]. Consequently, this method was believed to be a useful method in enhancing the redistricting process because the AHP allows decision makers to express their judgments of pairwise comparison in fuzzy ratios for indicating its importance in the aggregation procedure. In addition, the fuzzy ratio is able to avoid unbalanced scale of estimations and its ability to adequately handle the uncertainty and imprecision associated with the mapping of the decision makers' perception to a crisp number [4]. Fuzzy value from the SOR mentioned will be integrated into the model for the decision making process in the integrated model. Two of the compactness measurements can be taken into account according to the user's option. Indeed, compactness measurement to be used here can be customized to any other measures, which will be more relevant to different redistricting goals because the primary aim of this research is mainly to provide a metaphor to integrate FMCDM in GIS for shape-based redistricting by using the multiple compactness measurements.

The general shape-based redistricting decision problem usually consists of (a) a number of alternatives, which refer to the individual district, denoted as A_i $(i = 1, 2, \ldots, n)$, (b) a set of evaluation criteria C_j $(j = 1, 2, \ldots, m)$ which some of them may refer to the compactness measurements, (c) a qualitative or quantitative assessment x_{ij} $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m)$ (referred to as *performance ratings*) representing the performance of each alternative Ai with respect to each criterion C_j, leading to the determination of a decision matrix for the alternatives, and (d) a weighting vector (referred to as *criteria weights*) representing the relative importance of the evaluation criteria with respect to the overall objective of the problem.

Then, this research uses a FMCDM approach based on the synthesis of the following concepts, including (a) fuzzy set theory, (b) AHP, (c) α-cut concept, and (d) Decision maker (District Planner), DM's attitude towards risk. By using the fuzzy numbers defined, a fuzzy reciprocal judgment matrix for the decision matrix for m criteria and n alternatives is given as

$$X = \begin{bmatrix} x_{11}\ x_{12} \ldots x_{1m} \\ x_{12}\ x_{22} \ldots x_{2m} \\ \ldots\ \ldots\ \ldots\ \ldots \\ x_{n1}\ x_{n2} \ldots x_{nm} \end{bmatrix} \qquad (1.\ 2)$$

where x_{ij} represent the linguistic assessments of the performance rating of alternative $A_i (i = 1,2,...n)$ with respect to criterion $C_j (j = 1,2,...m)$. The decision matrix is to be given by the DM based on the term set being defined in Figure B.1.2.

The weighting vectors for the evaluation criteria can be given directly by the DM or obtained by using pairwise comparison of the AHP. The weighting vectors W of the linguistics term for the criteria is as follow:

$$W = (w_1, w_2, ..., w_m) \qquad (1.\ 3)$$

A fuzzy performance matrix representing the overall performance of all alternatives with respect to each criterion can therefore be obtained by multiplying the weighting vector by the decision matrix. The arithmetic operations on these fuzzy numbers are based on interval arithmetic.

$$Z = \begin{bmatrix} w_1 x_{11} + w_2 x_{12} + \ldots + w_n x_{1m} \\ w_1 x_{12} + w_2 x_{22} + \ldots + w_n x_{2m} \\ \ldots \quad \ldots \quad \ldots \quad \ldots \\ w_1 x_{n1} + w_2 x_{n2} + \ldots + w_n x_{nm} \end{bmatrix} (1.\ 4)$$

By using a α-cut

$$\forall \alpha \in [0,1],$$

$$\overline{A}_n = [a_1^n, a_3^n] = [(a_2 - a_1)\alpha + a_1, -(a_3 - a_2)\alpha + a_3]$$

On the performance matrix (1.3), an interval performance matrix can be derived as in (1.4), where $0 \le \alpha \le 1$. The value of α represents the DM's degree of confidence in his/her fuzzy assessments regarding alternative ratings and criteria weights. A larger α value indicates a more confident DM, meaning that the DM's assessments are closer to the most possible value a_2 of the triangular fuzzy numbers (a_1, a_2, a_3).

$$Z = \begin{bmatrix} [z_{1l}^{\alpha}, z_{1r}^{\alpha}] \\ [z_{2l}^{\alpha}, z_{2r}^{\alpha}] \\ \ldots \\ [z_{nl}^{\alpha}, z_{nr}^{\alpha}] \end{bmatrix} \qquad (1.\ 5)$$

Incorporated with the DM's attitude towards risk using an optimism index λ, an overall crisp performance matrix is calculated as in (1.5), where $z_{i\alpha}^{\lambda'} = \lambda z_{ir}^{\alpha} + (1-\lambda) z_{il}^{\alpha}, \lambda \in [0,1]$

$$Z_{\alpha}^{\lambda'} = \begin{bmatrix} z_{1\alpha}^{\lambda'} \\ z_{2\alpha}^{\lambda'} \\ \ldots \\ z_{n\alpha}^{\lambda'} \end{bmatrix} \qquad (1.\ 6)$$

In this enhanced/integrated shape based redistricting model, $\lambda = 1$, $\lambda = 0.5$, and $\lambda = 0$ are used to indicate that the DM involved has an optimistic, moderate, or pessimistic view respectively. An optimistic DM is apt to prefer higher values of his/her fuzzy assessments, while a pessimistic DM tends to favor lower values.

The evaluation criteria or compactness measurement index in this model is a relative measure and not an absolute measure. When the district planner compare two districts, they will not ask "How do these districts score?" but "Which district is more compact?". Therefore, to evaluate the similarity of relative compactness judgments between each pair of measures, normalization may need to be considered. Therefore, to facilitate the vector matching process, a normalization process in regard to each criterion is applied to (1.5) by using (1.6), resulting in a normalized performance matrix expressed as in (1.7).

$$z_{i\alpha}^{\lambda} = \frac{z_{i\alpha}^{\lambda'}}{\sqrt{\sum_{i=1}^{n} (z_{i\alpha}^{\lambda'})^2}} \qquad (1.7)$$

$$N_{\alpha}^{\lambda} = \begin{bmatrix} z_{1\alpha}^{\lambda} \\ z_{2\alpha}^{\lambda} \\ \cdots \\ z_{n\alpha}^{\lambda} \end{bmatrix} \qquad (1.8)$$

The values of N_{α}^{λ} indicate the degree of preference with respect to the alternatives for fixed α and λ, respectively where $\alpha \in [0,1], \lambda \in [0,1]$. Indeed, this value is the enhanced compactness index (ECI), which consider both of the compactness measurements or evaluation criteria earlier. Therefore, the larger the value, the more the preference of the alternative.

In summarizing the Fuzzy-AHP method for getting the enhanced compactness measurement index above, this research present the steps required for the algorithm in Figure B.1.4.

1. Formulate redistricting decision problem as multi-criteria problem
2. Identify the hierarchical structure of the problem
3. Define membership function for each criteria
4. Apply IF-THEN decision rules to assign a linguistic and numeric value of each criteria variable
5. Obtain decision matrix by fuzzy number as expressed in (1.1) using AHP method based on fuzzy number defined and Figure B.1.2
6. Obtain weighting vector for the criteria as expressed in (1.2) using AHP method based on fuzzy number defined and Figure B.1.3
7. Obtain fuzzy performance matrix (1.3) by multiplying the decision matrix obtained at step 4 by the weighting vector determined at step 5
8. Obtain interval performace matrix (1.4) by α-cuts on the performance matrix determined at step 6
9. Obtain crisp performace matix (1.5) by DM's attitude towards risk represented by an optimism index λ
10. Calculate normalized performace matrix (1.7) by (1.6)
11. Get the enhanced compactness index

Figure B.1.4. : Algorithm for the Fuzzy-AHP for this study

One of the significant improvements of the new proposed redistricting algorithm is the availability of the parameter to show the attitude to risk, α and level of confidence, λ of the district planner. These parameters act as observation parameter, which aim to analyze the decision-making behavior of the decision makers. Therefore, we have defined the decision makers based on their attitude to risk into three different groups that are the optimistic, moderate or pessimistic decision makers.

B.1.3.4 Optimization by Dynamic Programming

The redistricting problems are treated as similar to the knapsack problem, which is maximization or minimization of a value. For example, a thief robbing a safe finds it filled with N items of various size and value but bag has only limited capacity (M). In contrast, getting compact district plan means to calculate the best combination of individual district for all district size up to total district plan size. In other words, the district plan is a plan that consists of many districts with different shape. Therefore, the adopted implementation method in this research is Dynamic Programming(DP) which is commonly used to solve the knapsack problem. The method was chosen based on two main reasons. First, DP takes time as the horizon and calculates the least cost path in the interval. It is similar to the redistricting process which ask for optimal compact district, so DP can build a good model in a bottom-up technique to solve redistricting problems. It allows for the breaking up of all problems into a sequence of easier subproblems which are then evaluated by stages and has the power to determine the optimal solution by solving each of these stages optimality [11]. Second, redistricting decisions was usually accompanied with many complicated considerations, so these might be numerical constraints or some constraints which were the experience of the experts. These constraints are difficult to solve by nonlinear programming or other methods but can be easily incorporated and solved by the DP method. In addition, when integrate with FMCDM, these constraints can even be systematically analyzed and solved.

B.1.4.0 RESULTS AND ANALYSIS

The evaluation of the developed shape based redistricting algorithm using forest blocking application as an prototype focuses on the applications and limitations of incorporating an enhanced compactness index based on multiple compactness measurement into redistricting technique by using fuzzy multiple criteria decision making. Overall performance of the developed shape based redistricting algorithm is evaluated using statistical tests. The statistical tests were applied on different conditions in order to find out the performance of the developed shape based redistricting algorithm under different circumstances. These tests used for evaluation are referring to existing standard use to define a district plan. For example, a plan's compactness is defined as the mean compactness of its districts used in Iowa and in Michigan in United State for political redistricting (Altman, 1998). Besides, other research mentioned that its least compact district determines the compactness of a plan. Consequently, in this study, statistical test that include Mean, Mean Deviation, Maximum, Minimum and their Difference of compactness indices for its districts is used to determine the compactness evaluation of a District Plan.

There were three main areas of evaluations and tests being carried out on the redistricting algorithm: (a) performance evaluation and comparison on district plan produced with and without the enhanced redistricting algorithm, (b) result and analysis on the advance FMCDM component, (c) result and analysis with respect the different input variable like interval to calculate slope and also the restricted boundaries such as river and license boundary using prototype. From the evaluation process shown that the proposed method has the ability to consider multiple criteria to ensure the compactness is within optimality in producing district plan. Although there is a necessary to have a precise understand on the shape optimal rule or the linguistic terms' definition of the most compact district, the scheme able to produce a Enhanced Compactness Index in order to represent the performance of the district within the district plan. Besides, it is definitely going to give adequate reflection of the district planner toward risk and their confidence in their subjectivity assessment. In conclusion, the enhanced redistricting algorithm really shows ability to shape based redistricting scheme is going to give better performance to draw or redraw district plan with optimal compactness.

The applicability of the methods is proven effective. An Integrated Compactness Index had been successfully calculated and incorporated into redistricting technique that discussed. This compactness assessment index is more descriptive and able to incorporate with natural feelings of district planners. Natural feelings here may include their confidence and their attitude to risk. The shape compactness information has been modeled flexibly by utilizing the Fuzzy Multiple Criteria Decision Making approach like Fuzzy AHP. The Fuzzy AHP approach allows the integration of both Application Dependent and Application Independent factors to be considered in the redistricting application. The consideration of Application Dependent criteria is compulsory whereby the consideration of Application Independent criteria here is the more unique factor to ensure the optimality of shape compactness. Besides, the enhanced shape-based redistricting technique is able to cope with fuzziness in the compactness assessment index. The triangular fuzzy number used to define the decision matrix and weight matrix able to help to define the fuzziness of optimality of the compactness.

The integration of multiple compactness measurement methods in enhanced redistricting algorithm able to gather the strengths of particular method and at the same time to reduce or minimize the weaknesses or lacks of other method. Manipulating the weight matrix that represents the importance of a particular measurement accomplishes it. Then, these weight vectors will then contribute to the enhanced compactness index, which will incorporate into the redistricting process. A prototype was built based on the conceptual design of the redistricting algorithm. The prototype demonstrated the application of redistricting algorithm developed. It also allowed the testing and evaluation to be implemented. The prototype managed to demonstrate the concept of the integration and the incorporation in the redistricting algorithm. Later, overall performance of the developed redistricting algorithm is evaluated using statistics tests. The statistics tests were applied on different condition in order to find out the performance of the developed shape based redistricting algorithm under different circumstances. The approach developed clearly has its advantages. These advantages include (a) better modeling of the uncertainty and imprecision associated with the fuzzy triangular number, (b) Cognitively less demanding on the district planner, and (c) adequate reflection of the district planner's attitude towards risk and their degrees of confidence in their subjective assessment. Real experience in applying the approach in selecting the most appropriate district plan in the prototype has reinforced these findings. The incorporation of shape information sources has enhanced the redistricting techniques by utilizing more sources of information. As a result, the redistricting technique no longer based on the application dependent information only. The incorporation is proven to redraw district boundary effectively when there is more than one criterion.

B.1.5.0 CONCLUSION

This research has contributed to the improvement of redistricting technique by incorporation fuzzy multicriteria decision making for compactness measurement index. This research has identified and defined the shape information sources that are applicable to the redistricting technique. During the design stage, the shape information sources like the multiple compactness measurements were being modeled to meet the requirements and specifications defined and studied. These information sources have also been incorporated into the redistricting technique. The success of definition, modeling, and incorporation of the tertiary information also highlighted the applicability of Multiple Criteria Decision Making approach and Fuzzy Logic approach in redistricting technique. The research has successfully designed and developed a redistricting algorithm used to incorporate shape information into redistricting technique. The procedures for knowledge acquisition, preprocessing, analyzing the multiple criteria, and draws the district plan to the user has been defined. The overall performance of the prototype designed according to the integrated algorithms was tested and proven with a very significance improvement on the redistricting process from different aspect of testing.

REFERENCES

[1] Altman, Micah (1997a); What are the Judicially Manageable Standards for Redistricting? Evidence from History; California Institute of Technology; Pasadena, California.

[2] Altman, Micah (1998); Redistricting Principles and Democratic Representation; Ph.D. Thesis; California Institute of Technology; Pasadena, California.

[3] Fuller Robert and Carlsson Christer (1996); Fuzzy Multi Criteria Decision Making: Recent Developments; Fuzzy Sets and Systems, 78; 139-153.

[4] Deng H.(1999); Multicriteria Analysis with Fuzzy Pairwise Comparison; International Journal of Approximate Reasoning 21(3): 215-231.

[5] Herrera, F. and Verdegay, J.L. (1995); Fuzzy Sets and Operations Research Perspectives; Technical Report; Department of Computer Science and Artificial Intelligent; Granada, Spain.

[6] Skiena, Steven, 1997; The Algorithm Design Manual; Department of Computer Science State University of New York, Springer-Verlag, New York; URL:http://evo.apm.tuwien.ac.at/AlgDesignManual/BOOK/BOOK5.

[7] Zimmermann H.J. and Sebastian H.J. (1995); Intelligent System Design Support by Fuzzy Multi-Criteria Decision Making and/or Evolutionary Algorithms; Fuzzy Systems, International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium, Proceedings of 1995 IEEE International Conference, Vol 1; 367-374.

[8] Ravi, V. and Reddy, P.J. (1999); Ranking of Indian Coals via Fuzzy Multi Attribute Decision Making; Fuzzy Sets and Systems, 103; 369-377.

[9] Jacob Jen-Gwo Chen and Zesheng He (1997); *Using Analytic Hierarchy Process and Fuzzy Set Theory to Rate and Rank the Disability*; Fuzzy Sets and System, 88; 1-22.

[10] Pamela McCauley Bell and Heng Wang (1997); *Fuzzy Linear Regression Model for Assessing Risks for Cumulative Trauma Disorders*; Fuzzy Sets and Systems, 92; 317-340.

[11] Render, Barry & Ralph M.Stair (1995); *Quantitative Analysis for Management;* International Edition; Prentice Hall, London

BIOGRAPHY

Yin Chai WANG:

Deputy Dean & Head of Imaging and Spatial Information Systems Core Group, Faculty of Information Technology, Universiti Malaysia Sarawak

Research field: Image Processing and Spatial Data Handling, Automatic Digitizing of Maps and GIS, GIS for Health Care, Artificial Intelligent, and Virtual GIS

Chin Wei BONG :

Postgraduate student, Faculty of Information Technology, Universiti Malaysia Sarawak

B.2. TENCON 2001

IEEE REGION 10
International Conference on Electrical and Electronic Technology
TENCON 2001
Singapore* Phuket Island * Langkawi
19 – 22 August 2001

Compactness Measurement Using Fuzzy Multicriteria Decision Making For Redistricting

Wang Yin Chai, Bong Chin Wei

ABSTRACT

This paper presents a new method for compactness assessment in redistricting planning using Fuzzy Multicriteria Decision Making. An Enhanced Compactness Index (ECI) representing the overall plan with respect to each criterion is obtained by using triangular fuzzy number. The ECI is generated based on the synthesis of the concepts of fuzzy set theory, AHP, α-cuts concept and index of optimism of district planners to estimate the degree of satisfaction of the judgements on a district plan. The proposed method is more flexible, simple and comprehensive with easy computation and efficiency which facilitates its uses in compactness measurement in redistricting application like school redistricting, election boundary redistricting and others. A case study on forest blocking is presented to demonstrate its applicability in redistricting applications with respect to their redistricting goals and criteria.

Index Terms: Artificial Intelligence, Fuzzy System, and Geographical Information System

B.2.1 Introduction

The aim for redistricting or districting is to foster community and it begs the tough question: where to draw the lines for the district? It is tremendous important that it reacts on how to take control on the space of a particular region. Nonetheless, current redistricting method tends to produce odd and bizarre shape of districts, which often characterised as non-compact district plan. Compactness measurement is useful to measure the compactness and continuity of a district plan. However, choosing the compactness measurements is a special challenge because there are over thirty distinct measures of compactness. Thus, this paper aims to enhance existing compactness measurements and use it in redistricting algorithm to improve the shape of the district plan.

Subsequently, this paper discusses on the use of Fuzzy Multicriteria Decision Making (FMCDM) on compactness measurement in redistricting process. In simple, the aims of this paper is as the following:

1) To produce better shape assessment index according to integration of multiple compactness measurement method that is more descriptive and able to incorporate with natural feelings of district planners with fuzziness.
2) To incorporate the new index into redistricting to generate an optimal compact district.
3) To consider the restricted boundary like river, political boundary, and others during redistricting.
4) To work and perform in an environment that is able to manage the spatial and non-spatial data, their relationship and dependency.

Fig B.2.1 briefly shows the revised enhanced redistricting model in terms of the input, process and output. Thus, the input data are both the spatial data and the related non-spatial data. The domain of the proposed model is on enhancing the decision support system in maintaining the compactness and continuity of redistricting to produce the optimal compact district plan with FMCDM. Meanwhile, GIS technology acts as a supportive tool because of its capability to handle the relationships of spatial data between non-spatial or namely attributes data.

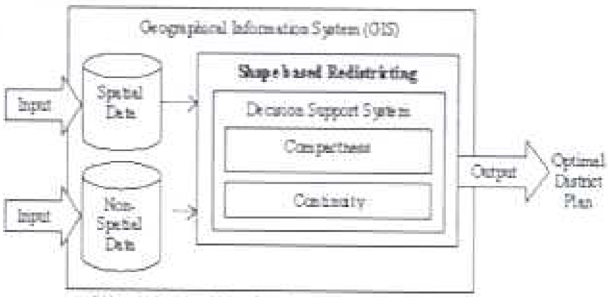


Fig B.2.1: Revised redistricting model

B.2.2. Architecture Design

This paper applies the concept of FMCDM at each redistricting stage shown in the semantic diagram in Fig B.2.2. Shape Optimal Rules (SOR) is the data store, which consists of the knowledge acquisition from the district planners. This data store includes the objective and the fuzzy rules sets for the FMCDM. Two other data stores: Application Dependent Data Store (ADD) and Application Independent Data Store (AID) provide the necessary input data for the algorithm. The initiate stage is in Data Preparation Module (DPM), followed by three stages (S_1, S_2, and S3) for creating district occurs and decisions are made with FMCDM. Lastly, Combine Optimal District Module (COD) is used to get the final optimal district plan.

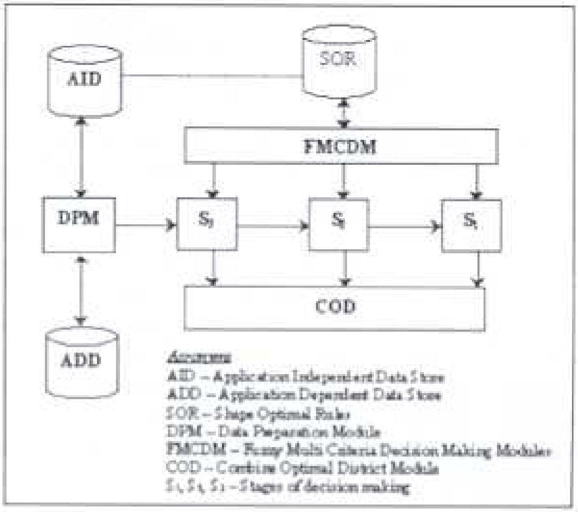


Fig B.2.2: Semantic Diagram of the enhanced model

B.2.3. Data Preparation for the Input

Firstly, we prepare and group the data from the data store into smallest features to represent the redistricting criteria. This input data format is in polygon and point features because it links to an attribute record. The paper determines to use triangles as the input polygon features because triangulation is the mother of all polygon partitioning problems triangulation, and the interior of all kind of polygons can be completely partitioned into triangles [1]. After that, the redistricting area will be intersected with a license boundary, which provides the exact location of the redistricting process.

B.2.3.1.*Shape Optimal Rules*

Shape Optimal Rules is an important role to determine the optimality of district compactness. The compactness measurement methods give a gray area (Fig B.2.3) for district planners on the district shape assessment and we use the theory to formalize the gray area to specifically express them. Fuzzy set theory provides fuzzy numbers that are easy to use in expressing qualitative assessments of the Decision-Makers, DM [2,3,4,5]. Therefore, we define the necessary rulesets for the FMCDM process in order to generate the Enhanced Compactness Index.

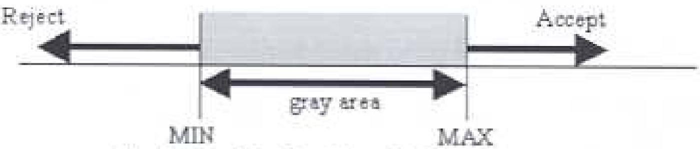


Fig B.2.3: The Decision Making Process

Specifically, we determine two different compactness measurement methods to consider the compactness criteria and they act as two different separate criteria. They include the non-Euclidean measurement based on fractal dimension and the Euclidean measurement based on area-perimeter ratio. The higher the fractal dimension, the more complex it is for the district boundary. Subsequently, the district compactness is determined by the value on the fractal dimension. Thus, we choose to use box counting dimension to calculate the fractal dimension. First, we need to cover an image by 'r'-size boxes and determining how many boxes of a particular size 'r' intersect the image. Thus, the number of boxes of size 'r' is needed to cover the image is given by:

$$N(r) = \frac{l}{r^{D}} \quad (20)$$

log(box count) = a + b log(box size) (21)

To estimate the box-counting dimension, the Euclidean space containing the image are divided into a grid of boxes of size 'r' and counting such boxes N(r) which are non-empty. Then, the size 'r' is changed to progressively smaller sizes and the corresponding numbers of non-empty boxes are counted N(r). The logarithm of N(r) versus the logarithm of 1/r gives a line whose gradient corresponds to the box dimension. The sequence of mesh sizes for grids is usually reduced by a factor of 1/2 from one grid to the next. Therefore, if the number of boxes counted increased by a factor of when the box size is halved, then the fractal dimension is equal to D [6]. The box sizes and number of boxes intersected for selected district boundary is converted to logarithm format. When the data are plotted on a log-log graph based on the results where box count is the number of boxes overlapping the feature, and box size is the length of one side of the box. The intercept is represented by a, and b is the slope. The fractal dimension D is represented by the absolute value of the slope b.

Another Euclidean compactness measurement is used to control the inconsistency of the fractal dimension. The Euclidean measure for the compactness is as below:

$$compactness = \frac{Perimeter^{2}}{Area} \quad (3)$$

A preliminary experiment is conducted and the result proves the consistency of this method comparing to the fractal dimension in the non-Euclidean measurement. Both of these measurements are used together in the case study to provide the ECI value for the use in the FMCDM in the next section.

Then, we define the rules by linguistic terms and their membership function. Linguistics terms are defined for each of the selected criteria and their weighing vector. In brief, the weighing vector represents the decisions of the district planners at the importance on each criterion among all the criteria. We apply the fuzzy set theory in these weighing vectors for the ranking of the relative importance's during the decision making process. In addition, we use the definition of the interval of confidence at level α (α-cuts). This helps characterize the triangular fuzzy number as

$$\forall \alpha \in [0,1],$$
$$\overline{A}_{\alpha} = [a_1^{\alpha}, a_3^{\alpha}] = [(a_2 - a_1)\alpha + a_1, -(a_3 - a_2)\alpha + a_3] \quad (4)$$

DM may often meet the situation where it is difficult for them to choose or reject alternatives. Thus, only the YES/NO method needs to be improved due to the subjective judgments of the DM [7]. One alternative method is for the DM to give an effectiveness level (x) to each of the selected criteria (C). The effectiveness levels belong to a set of linguistic terms that contains various degrees of preference required by the DM. Then, this study models the subjectiveness and vagueness of the decision making process by linguistic terms which has been found intuitively easy to use [2]. This study also use linguistic terms μ (effectiveness, x) = {Very Poor (VP), Poor (P), Fair (F), Good (G), Very Good (G)}, defined in Fig 4. A membership function, which assigns to each linguistic term with a grade of membership, is associated with the fuzzy set in [0, 1]. When the grade of membership for a linguistic term is one, it means that the linguistic term is absolutely in that set and verse vice. Borderline cases are assigned values between zero and one.

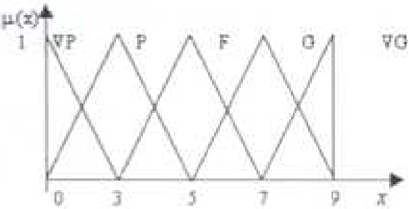


Fig B.2.4: Membership functions, $\mu(x)$ and the linguistics terms for each effectiveness (x) of each criteria (C)

The different weights to the criteria are determined by pairwise comparison of the relative importance of criteria, the pairwise comparison matrix $E=[e_{ij}]_{t\times t}$ is established, where e_{ij} represents the quantified judgments on pairs of criteria C_i, and C_j. The comparison scale ranges from 1 to 9, each representing the concepts of: 1: equally important; 3: weakly more important; 5: strongly more important; 7: demonstratively more important; 9: more important; 2, 4, 6, and 8 are intermediate values between adjacent judgments. For example, $e_{ij} = 5$ means C_i is strongly more important than C_j. We also facilitate the making of pairwise comparison with triangular fuzzy numbers defined in Table B.2.1 are used. A triangular fuzzy number $\bar{x}$ expresses the meaning of 'about x', where $1 \leq x \leq 9$, with its membership function. Fuzzy number $\bar{x}$ used by [2] is revised to better reflect the decision situation involved.

Table B.2.1: Linguistic terms used by the decision matrix

Linguistic term	Very Poor (VP)	Poor (P)	Fair (F)	Good (G)	Very Good (VG)
Fuzzy number	$\bar{1}$		$\bar{x}$		$\bar{9}$
Membership function	(1, 1, 3)		$(x-2, x, x+2)\ for\ x=3,5,7$		(7, 9, 9)

B.2.3.2. FMCDM by using Fuzzy-AHP

A FMCDM method that called Fuzzy Analytical Hierarchy Process or Fuzzy AHP is selected in this paper. Analytical Hierarchy Process is a multi criteria method, which uses hierarchic structures to represent a decision problem, and then develops priorities for the factors based on DM's judgement. Fuzzy AHP method was discussed and being used for ranking of Indian Coals in industrial use [8], rate and ranking the disability [9], assessing risk of cumulative trauma disorders) [10], evaluating weapon system [4] and many other applications.

Therefore, the redistricting decision problem consists of (a) a number of alternatives, which refer to the individual district, denoted as A_i $(i = 1, 2, \ldots, n)$, (b) a set of evaluation criteria C_j $(j = 1, 2, \ldots, m)$ which some of them may refer to the compactness measurements, (c) a qualitative assessment x_{ij} $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m)$ representing the performance of each alternative Ai with respect to each criterion C_j, leading to the determination of a decision matrix for the alternatives, and (d) a weighting vector representing the relative importance of the evaluation criteria with respect to the overall objective of the problem.

Then, this study uses a FMCDM approach based on the synthesis of the following concepts, including (a) fuzzy set theory, (b) AHP, (c) α-cut concept on the level of confidence, and (d) Decision maker (District Planner), DM's attitude towards risk. Consequently, the subjectivity and imprecision of the evaluation process are adequately handled, and the complex and unreliable process of the ranking procedure starts at the determination of the criteria importance and alternative performance. By using the fuzzy numbers defined in Table B.2.1, a fuzzy reciprocal judgement matrix for the decision matrix for m criteria and n alternatives is given as

$$X = \begin{bmatrix} x_{11}\ x_{12} \dots x_{1m} \\ x_{12}\ x_{22} \dots x_{2m} \\ \dots\ \dots\ \dots\ \dots \\ x_{n1}\ x_{n2} \dots x_{nm} \end{bmatrix} \quad (5) \qquad W = (w_1, w_2, \dots, w_m) \quad (6)$$

where x_{ij} represent the linguistic assessments of the performance rating of alternative $A_i (i = 1,2,..n)$ with respect to criterion $C_j (j = 1,2,..m)$. The weighting vectors for the evaluation criteria is given directly by the DM or obtained by using pairwise comparison of the AHP as defined in Table 1. Then, the model obtains a fuzzy performance matrix (8) representing the overall performance of all alternatives, with respect to each criterion by multiplying the weighting vector with the decision matrix. The arithmetic operations on these fuzzy numbers are based on interval arithmetic.

$$Z = \begin{bmatrix} w_1 x_{11} + w_2 x_{12} + \dots + w_m x_{1m} \\ w_1 x_{12} + w_2 x_{22} + \dots + w_m x_{2m} \\ \dots \quad \dots \quad \dots \quad \dots \\ w_1 x_{n1} + w_2 x_{n2} + \dots + w_m x_{nm} \end{bmatrix} \quad (7)$$

By using a α-cut (4) on the performance matrix (7), an interval performance matrix can be derived as in (8), where $0 \leq \alpha \leq 1$. The value of α represents the DM's degree of confidence in his/her fuzzy assessments regarding alternative ratings and criteria weights. A larger α value indicates a more confident DM. Incorporated with the DM's attitude towards risk using an optimism index λ, an overall crisp performance matrix is calculated in (9) by $z_{i\alpha}^{\lambda'} = \lambda z_{ir}^{\alpha} + (1-\lambda) z_{il}^{\alpha}, \lambda \in [0,1]$

$$Z = \begin{bmatrix} [z_{1l}^{\alpha}, z_{1r}^{\alpha}] \\ [z_{2l}^{\alpha}, z_{2r}^{\alpha}] \\ \dots \\ [z_{nl}^{\alpha}, z_{nr}^{\alpha}] \end{bmatrix} \quad (8) \qquad Z_{\alpha}^{\lambda'} = \begin{bmatrix} z_{1\alpha}^{\lambda'} \\ z_{2\alpha}^{\lambda'} \\ \dots \\ z_{n\alpha}^{\lambda'} \end{bmatrix} \quad (9)$$

λ = 1, λ =0.5, and λ = 0 are used to indicate that the DM involved has an optimistic, moderate, or pessimistic view respectively. An optimistic DM is apt to prefer higher values of his/her fuzzy assessments, while a pessimistic DM tends to favor lower values. After the facilitation of the vector matching process, a normalization process in regard to each criterion is applied to (10), resulting in a normalized performance matrix expressed as in (11).

$$z_{i\alpha}^{\lambda} = \frac{z_{i\alpha}^{\lambda'}}{\sqrt{\sum_{i=1}^{n} (z_{i\alpha}^{\lambda'})^2}} \quad (10) \qquad N_{\alpha}^{\lambda} = \begin{bmatrix} z_{1\alpha}^{\lambda} \\ z_{2\alpha}^{\lambda} \\ \dots \\ z_{n\alpha}^{\lambda} \end{bmatrix} \quad (11)$$

The values of N_{α}^{λ} indicate the degree of preference with respect to the alternatives for fixed α and λ, respectively where $\alpha \in [0,1], \lambda \in [0,1]$. Indeed, this value is the Enhanced Compactness Index (ECI), which considers both of the compactness measurements or evaluation criteria earlier. Therefore, the larger the value, the more the preference of the alternative. In summarizing the

Fuzzy-AHP method for getting the Enhanced Compactness Index above, this paper presents the steps required for the algorithm in following:

1. Formulate redistricting problem as multicriteria problem
2. Identify the hierarchical structure of the problem
3. Define membership function for each criteria
4. Apply IF-THEN decision rules to assign a linguistic and numeric value of each criteria variable
5. Obtain decision matrix by fuzzy number in (5) using AHP method based on fuzzy number in Table B.2.1and Fig B.2.4
6. Obtain weighting vector for the criteria in (6) using AHP method based on fuzzy number in Table B.2.1 and Fig B.2.4
7. Obtain fuzzy performance matrix (7) by multiplying the decision matrix obtained at step 5 by the weighting vector determined at step 6
8. Obtain interval performace matrix (8) by α-cuts on the performance matrix determined at step 7
9. Obtain crisp performace matix (9) by DM's attitude towards risk represented by an optimism indexing, λ
10. Calculate normalized performace matrix (11) by (10)
11. Get the Enhanced Compactness Index (ECI)

Later, the problem is divided into sub-problems of the same kind called stages. These stages are the logical sub-problems, which consists of input variables or state variables those are the possible beginning situations or conditions of a stage. The next step (at Stage 1, S_1) is to solve the last or the worst stage of the problem for all possible conditions or states using an array or sequences of arrays. Solving the last or worst stage is trivial and it acts as the preprocessing period to provide data to the decision scheme in FMCDM. Then, the scheme working backward from the last stage and solve each intemediate stage (at Stage 2, S_2 and Stage 3, S_3). The decision to make at every stage aims to reach the optimal compact. In general, the notation for the important concepts are as below:

$$S_n = \text{Input to stage } n \qquad (12)$$

$$D_n = \text{Decision at stage } n \qquad (13)$$

$$R_n = \text{Return at stage } n \qquad (14)$$

$$S_{n-1} = \text{Output from stage } n \qquad (15)$$

At each stage, the model performs the redistricting process by annexes adjacent triangular to create district after the system sort and select the polygon. The process continues with checking and merging the smaller polygons until it reaches a polygon with the ideal area size. The transformation function for stage 2 (t_2), convert the input to stage 2 (s_2), and the decision made at stage 2 (d_2) is to obtain the output from stage 2 (s_1). The transformation process depends on the input and decision at any stage, it can be represent as t_2 (s_2, d_2). Every return from the stage n in the previous module will be sent to this module.

B.2.4. Prototyping Developing

We select a case study to verify the practically of the proposed model. The selected case study for implementation is on Forestland Blocking for dividing the forestland into different blocks according to specific criteria for the purpose of enrichment planting and industry planting. Therefore, the input for the prototype including the elevation for contour as the application dependent factors to draw the district line. The physical environment of the system prototype is within GIS software called ARC\INFO 7.2.1 from ESRI, which provides most of the spatial function needed for the prototype development process. Then, the basic hardware environment includes Window NT 4.0 with 32 MB RAM and Hard Disk 2.0 GB. The language used to implement the model is Arc Macro Language (AML), which is the scripting language, used in the ARC\INFO.

B.2.4.1.Knowledge Acquisition

An experiment is conducted in order to find the performance of the selected criteria which regarding to compactness. Thus, the experiment is conducted toward the criteria based on fractal dimension and based on area-perimeter ratio. This experiment is extremely important to obtain the interval value in x-axis of the membership function or the effectiveness level, x for each linguistic term as mentioned earlier. A set of 29 special polygons sample as in Fig B.2.5 is created to survey on both of the compactness index produced by the fractal dimension and the Euclidean measure by area-perimeter ratio.

The criteria considered during the redistricting process for the case study is summarized in the Fig B.2.6 and it is relying on the district area size, fractal dimension that represent the district boundary complexity and the Euclidean measurement to represent the district compactness and continuity. Let assume the ideal area is A and optimal range of district area is within B meter square. Therefore, the paper defines as "Very Good (VG)", "Good (G)", "Fair(F)", "Poor(P)" and "Very Poor(VP)" for every quarter changes of B size as illustrated in Fig B.2.7. The linguistic terms for district compactness index for EM is defined as {"Very Very Compact (VVC)", "Very Compact (VC)", "Compact (C)", "Not Very Compact (NVC)", "Not Compact (NC)"} as shown in Fig B.2.8. Meanwhile, the definition of FD criteria as {"Good (G)", "Quite Good (QG)", "Average (A)", "Quite Complex (QC)" and "Very Complex. (VC)"} as shown in Fig B.2.9.

Fig B.2.5: Sample set of district shape

Table B.2.2: Result for two different measure in Fig B.2.5

District Number	*Euclidean Measure*	*Fractal Dimension*
2	43.3004	1.0126
3	14.0527	1.0102
4	12.5779	0.9841
5	16.0017	1.1261
6	45.6757	0.9893
7	40.9493	1.0043
8	16.5371	1.0089
9	19.9991	1.0562
10	20.5611	1.0155
11	30.7006	1.0024
12	72.6556	1.0127

13	40.8622	0.9918
14	19.7707	1.0084
15	16.0162	0.9949
16	113.3842	1.0278
17	23.6295	1.0102
18	39.8167	1.0070
19	28.0176	0.9716
20	15.5493	0.9913
21	19.4320	0.8957
22	55.2553	0.9977
23	18.6523	1.0265
24	21.1447	1.0279
25	37.9112	1.0515
26	21.5037	1.0800
27	18.9608	1.0048
28	43.7079	1.0126
29	18.8888	1.0312

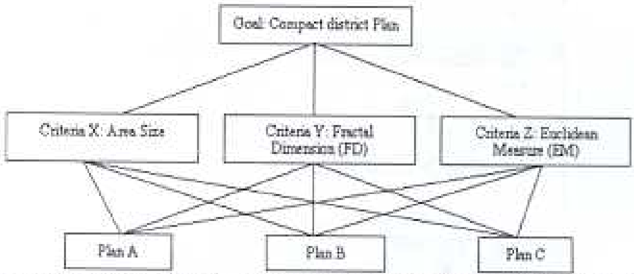


Fig B.2.6: Three criteria are considered in the problem definition

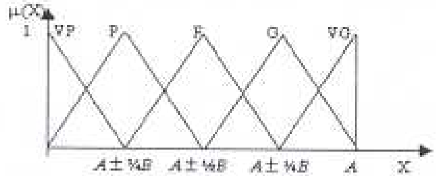


Fig B.2.7: Membership functions, μ(X) and the linguistics terms for ideal district size, A

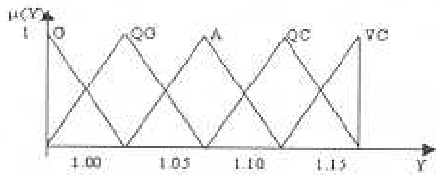


Fig B.2.8: Membership functions, μ(Y) and the linguistics terms for Fractal Dimension, Y

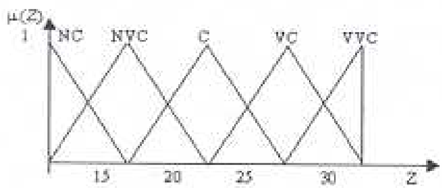


Fig B.2.9: Membership functions, μ(Z) and the linguistics terms for Euclidean Measures, Z

B.2.4.2. Prototype Design

There are four main modules as illustrated with Module One to Four in the Diagram 0 of the Data Flow Diagram in Fig B.2.10. These modules are specifically named as Initialization module, Districts Conquering module with Dynamic Programming technique, Fuzzy Multicriteria Decision Making (FMCDM) module and Combine Optimal District (COD) module. Initiation module is to prepare the input data meanwhile the district conquering module will select and conquer the district plan with dynamic programming method. Then, decisions will be made at the FMCDM module while final result will be integrated in COD module. Besides, there are four main data stores to store the contour layer, license layer, river layer, and the shape optimal rulesets.

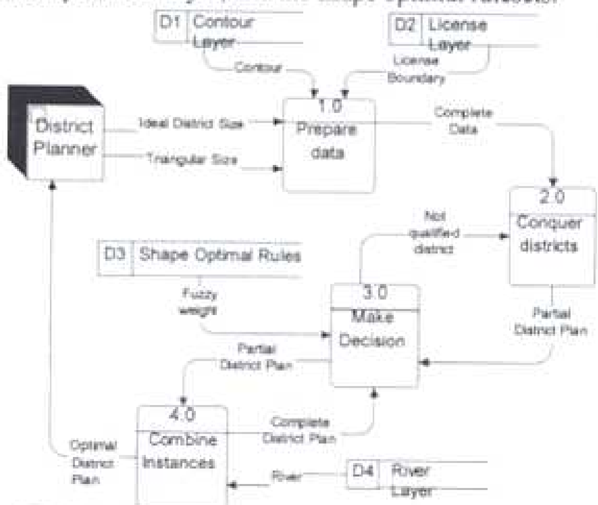


Fig B.2.10: Diagram 0 of the Data Flow Diagram

The main input data needed are the data source for contour, license boundary, and river. Besides, the district ideal size and triangles sizes are required too for the Initialization module and District Conquering module. Besides, for the FMCDM module, weight values for EM, FD, district area size, degree of confidence and attitude to risk are needed too. The necessary spatial data for the paper case study are digitized topographic map with 1:50,000 scale. Contour with 100 feet interval is the basic requirement to provide the slope data. Slope in degree acts as the application dependent data requirement or criteria for the redistricting process in the prototype development.

B.2.5. Result and Analysis

This study defines the DM based on their attitude to risk into optimistic, moderate or pessimistic DM. For each group of the DM, the paper let the level of confidence, $\alpha = 0.1, 0.3, 0.5, 0.7$ and 0.9, the paper can obtain the Enhanced Compactness Index for alternatives of district plan based different combination of weight values. The results in Fig B.2.12 to Fig B.2.14 show that district plan based on weight value 177 for the three selected criteria, are clearly the best choice under almost any degree of confidence of the DM with various attitudes towards risk. Besides, we also conduct the evaluation on district plans with and without the enhanced algorithm. The district plan generated uses triangles size = 5,000; ideal district size = 1,000,000; weighing vector for District Area Size, Fractal Dimension, Euclidean measure on area-perimeter ration = triplet 717; confidence level = 0.5; attitude to risk =0.5. The results from this analysis show that the enhanced algorithm produces more compact district plan (Fig B.2.11). Besides, the district plan without enhanced algorithm is odd or bizarre. In short, the district boundary for the district plan without enhanced algorithm is much more complex. In other word, the district is less compact.

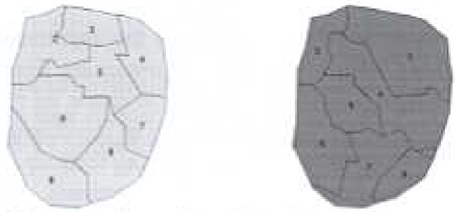


Fig B.2.11: District plan with and without enhanced algorithm

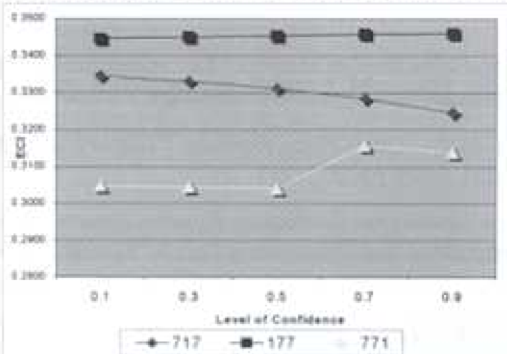


Fig B.2.12: Mean of ECI for Optimistic

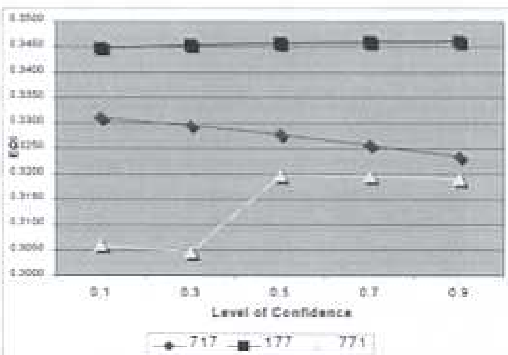


Fig B.2.13: Mean ECI of Moderate

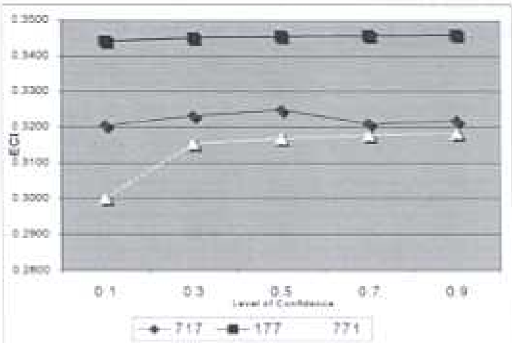


Fig B.2.14: Mean ECI of Pessimistic

B.2.6. Conclusion

The paper has been completed successfully despite all the constraints being posed to. The most profound accomplishment among all is the success in obtaining Enhanced Compactness Indexing and incorporating it into the Redistricting technique to enhance its result. Though the information sources are limited, it is enough to prove the applicability of the developed redistricting algorithm to incorporate shape compactness information into redistricting technique. The algorithm is simple and yet efficient as compared to other more complicated methods. The result is satisfactory and comparatively better than other traditional approaches as well.

REFERENCES

[1] Skiena, Steven, "*The Algorithm Design Manual*," Department of Computer Science, State University of New York, Springer-Verlag, New York, 1997.

[2] Deng H., "*Multicriteria Analysis with Fuzzy Pairwise Comparison*," International Journal of Approximate Reasoning 21(3), 1999, pp. 215-231.

[3] Chen Shyi-Ming, "*A new Method for Evaluating Weapon System Using Fuzzy Set Theory*;" IEEE Transaction on Systems, Man and Cybernetics Part A: System and Human, Vol 26, 1996, pp. 493-496.

[4] Mon, Don-Lin, "Evaluating Weapon System Using Fuzzy Analytical Hierarchy Process Based On Entropy Weight," IEEE, 1995, pp. 598-591.

[5] Yeh C.H. and Deng H., "An Algorithm for Fuzzy Multi-Criteria Decision Making," Proceedings of the IEEE First International Conference on Intelligent Processing Systems (ICIPS'97), 1997, pp. 1564-1568.

[6] Knight , Jeffrey L., "*GIS Based Compactness Measurement Using Fractal Analysis,*" George Mason University Fairfax, Virginia, 1997.

[7] Ron Kwok et. al., "*Improving Group Decision Making : A fuzzy GSS Approach,*" Dept of Information System, City University of Hong Kong; Tat Chee Avenue, Kowloon, Hong Kong, 1998.

[8] Ravi, V. and Reddy, P.J., "Ranking of Indian Coals via Fuzzy Multi Attribute Decision Making," Fuzzy Sets and Systems, 103, 1999, pp. 369-377.

[9] Jacob Jen-Gwo Chen and Zesheng He, "*Using Analytic Hierarchy Process and Fuzzy Set Theory to Rate and Rank the Disability*," Fuzzy Sets and System, 88, 1997, pp. 1-22.

[10] Pamela McCauley Bell and Heng Wang, "*Fuzzy Linear Regression Model for Assessing Risks for Cumulative Trauma Disorders*", Fuzzy Sets and Systems, 92, 1997, pp. 317-340.

Conference on Information Technology in Asia 1999: CITA '99
The Asian Regional Conference of IFIP WG 9.4
16-17 September 1999, Kuching, Sarawak

State of Art in Geographical Based Redistricting

Bong Chin Wei, Dr Wang Yin Chai

Abstract

Geographical features on redistricting applications are issues in political geography. However, to draw district line on the map is not only applicable on political purpose only, it also plays an important role for school, business and forests planning related system. Moreover, redistricting is tremendously important that it reacts on how government takes control on the space of a particular region. However, redistricting is a complex system and requires an interdisciplinary approach to its solution (Altman, 1998). Spatial redistricting measurement criteria include geographical size, shape and lines. These bring to issues of compactness and continuity. Many standard and rules have been introduced and argued but there is no any standard that being established. There are many problems occurred in the implementation stage and many methods or algorithms developments unable to produce the optimum or best result. Altman (1998): Many scholars have proposed ways to measure compactness but few have systematically analyzed these measures. Ill-compactness is a warning signal that requires justification, or that compactness is a useful, neutral, and objective criterion for redraw a fair district line. GIS will be an effective tool in helping redistricting process. In addition, fractal analysis in GIS environment may provide a clear approach on analyzing the spatial structure and form than current compactness measures.

B.3.1. Introduction

Geographical based redistricting is increasingly important because geographical criteria can settle criteria like unbiasedness and competitiveness (Cain, 1984). In many countries, the electoral laws specify that geographical, or certain geographical factors, be taken into account when delimiting electoral district lines. In redistricting application, geographic criteria are criteria related to geographical boundaries, size and/or shape (Handley, 1997; Altman, 199. A boundary authority may be asked to consider factors from either or both criteria. However, the concept of shape is often difficult to express (Knight, 1997). This difficulty manifest itself in the evaluation of shape in geographic study.

Automated redistricting may be intractable as it is a large mathematical problem and this problem of finding an optimal districting plan is computationally complex (Altman, 1998). The redistricting problem of finding an optimal districting plan is computationally difficult that it is unlikely that any minor increase in speed of computers will enable to solve it. Secondly, Altman (1998) argues that although there are more than thirty different measures, few have systematically analyzed these measures. Current methods are inadequate because purely automated redistricting techniques remain unsatisfactory.

Current redistricting algorithms or formula have limited functions to satisfy spatial needs. The natural geography of earth that comprises of forests, rivers, hills and mountains creates a situation that make the process of creating the boundaries more difficult especially in the rural areas where administrative transportation facilities are insufficient (Ngu, 1998). Therefore, implementation for redistricting a boundary based on the geographical compactness and continuity seems to have great limitation because it causes the redistricting process even more complex and intractable. However, conventional programming method for spatial redistricting algorithm has limitation on the implementation process. Spatial complexity for redistricting like topography factors causes inflexibility on conventional computer (Altman, 1998)

B.3.2 Redistricting Application and Their Importance

Redistricting is a term that always refers in election boundary application. However, redistricting process is not only relevant to election system, but it is needed in school, law enforcement and even forests planning related systems. Redistricting is tremendously important that reacts on how a government takes control on the space of a particular region. The spaces within which political, social, cultural, and economic processes unfold are not simply static backdrops or locations referents for human events, but the products of distinct territorial structures, identities, and ambitions, and they are deeply implicated in social and political change (Agnew, 1994). Therefore, redistricting process is important to redraw the lines of a district based on the social and natural changes from time to time.

In the process of redistricting election boundary, the more direct control the parties have over the drawing of district lines, the more they are able to manipulate these lines to produce the results they want (Hearne, 1999). Liittschwager (1997) explains that political districts not only allow state and local governments to provide citizens with fire departments, schools, and law enforcement, but determine who can vote for whom in local, state, and federal elections. In Iowa and elsewhere, state lay and court opinions require that political districts have nearly equal populations and be as contiguous and compact as possible (Liittschwager, 1997). Crowded schools and low-test scores are problems in educational institution in some countries today (Jones, 1998). Therefore, redistricting is now being considered in order to reassign pupils to the new school to reduce enrollments in schools that currently operating over its capacity. Redistricting will allow more children to go to schools closest to their houses and it will help in cutting transportation costs (Newsom, 1997). Then, it is helping in cutting transportation process costs. In short, redistricting is not only important for political district, but also for school and forest district. Moreover, business district also uses redistricting process to have sales territory management. It re-aligns sales personnel to more efficient sales areas. By then, customer service is improved and revenue per coverage area will be increased when sales territories are optimized. In addition, improved sales territories will result in timesaving and allow sales representatives to follow up more new business opportunities.

B.3.3 Geographical Concerns In Redistricting

The important factors related to the geometric shape of a district include continuity and compactness. For instance, judicial court in United States holds that districts should not be oddly shaped and that all pieces of land in a district should be inter-connected (Handley & Maley, 1997; Knight, 1997). Although the shape of these districts is not the basis for several of court's decision in most of the country, the compact districts are important in creating the district boundaries. Therefore, the research will discuss the meaning and definition of these criteria in the following section in order to clarify the important terms in this research.

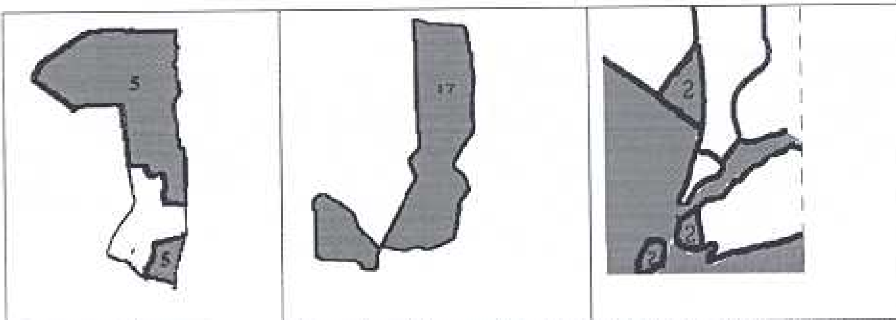
B.3.3.1 Compactness

Compactness is defined as how tightly a shape is "packed", and it is often used as a characteristic to describe shape (Shiode, 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). Then, Knight (1997) describes this includes a mathematical analysis of the shape, regularity of the spatial shape, and how regularly the population is distributed within the district. Therefore, irregular geographical boundaries and significant land areas may justify unequal district lines if such district lines follow a significant geographical feature or political subdivision boundary (Altman, 1997a).

Compactness may be determined by an analysis of the function 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). Thereafter, different applications will treat compactness with various meanings. For examples, compactness refers to population dispersion in political redistricting, but it refers to student distribution in school redistricting. However, the basic definition for compactness that will be used in this research is referring to geographical compactness which regarding to the shape of the district.

B.3.3.2 Continuity

Secondly, continuity is another popular geographic principle in redistricting. Many state constitutions in the United States list continuity as a requirement for legislative districts (Handley & Maley, 1997). The requirement is to ensure all districts to be inter-connected. Altman, (1997b) divides districts into three categories in order of divergence from real world continuity: practically contiguous, questionably contiguous, and non-contiguous. As in figure below, practical contiguous are all districts those are formally contiguous or depart from continuity because of islands off the coast of the district. Questionably contiguous is second category districts that are otherwise contiguous but that contained islands that are not directly off the coast of the district, districts that are non-contiguous but could be connected by straight bridges, and districts that are connected only by "points" (Altman, 1998). Districts which are non-contiguous are non compact.



Non-contiguous	Questionable contiguous (it is connected only at a single point)	Questionably contiguous (the island portions of the district are not joined to the nearest mainland district)

Figure B.3.1: Three different districts with their continuity category.

B.3.4 Conventional Compactness Measures

Recent surveys of compactness criteria list more than thirty different measurement formulas (Altman, 1998). Table B.3.1 shows the different type of compactness measures.

Table B.3.1: Compactness Measure (Altman, 1998)

Measurement	Comments
Area/perimeter	A/0.282P
Normalized Area/Perimeter (Norm)	$A/(0.282P)^2$
Area of Circle (AC)	The ratio of the district area to area of minimum circumscribing (Normalized to the [0,1] interval)
Length/Width (LW)	The length of the minor axes/major axes for the best fitting ellipse

As in Table B.3.1, geographical features that are taken into account in conventional redistricting are including the area, perimeter, length and width of the districts' shape. There are numerous techniques to measure these elements. For examples, Manley (1995) used pixel crack method and boundary pixel method to find perimeter. Pixel Crack counting returns the number of neighbors those are not the same labels using 4-neighbor connectivity. Therefore, the perimeter equals the sum of the 4-neighbors with different labels for all the pixels in the component that have any of their four connected neighbors in a different component. Thus, compactness measures the amount of area per amount of perimeter. The more compact a component is the lower the value of compactness. Therefore, a circle is the most compact object.

However, complexity and unsatisfactory of compactness measures in term of their shape equality has caused the result of odd or bizarre district shape. Therefore, Altman (1998) uses an axiomatic analysis to join formal measures of compactness to analyze the consistency of these different compactness criteria. These definitions can help now formally define what it means for a compact measure to capture shape: Let a shape $S = \{ s_1 ... s_i \}$ be a finite, nonempty set of simple, continuous, closed, non-overlapping subsets of the plane where *Area* $(s_i \cap s_j)$=*Perimeter* $(s_i \cap s_j)=0, \forall i \neq j$ (refer Figure 3.2). Let $P{:}S \rightarrow \Re_+$ be length of the perimeter of the shape, and let $A:S \rightarrow \Re_+$ be the area of the shape. Let a compactness measure C, be a function $C:S \rightarrow \Re$.

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. The research can 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:

5. *Minimal dissection*: Any two shapes with identical convex hulls, S_1 has a strictly smaller area, then S_1 should be judged 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 less compact.

Unfortunately, from the evaluation of Altman's axioms, most of the conventional redistricting measures show inconsistencies. Most compactness indices reflect at least one principles of shape manipulation but no others. In most cases, these measures obviously satisfy one shape axiom, but violate others. The shapes are not generally rectangular, straight or circular and cannot be represented by a simple analytical function. They have a more complex nature that is not easily defined. The dilemma can be traced to the problem of unambiguously describing the shape of an object. The seemingly simple task of describing a complex shape often is a intimidating task.

B.3.5 GIS Based Compactness Measures Using Fractal Analysis

Important development in the late 1980s contributed to a far greater use of computer for redistricting in the 1990s were the development of geographical information system(GIS) software. GIS software is an integrated system for displaying demographic data on the computerized map. In 1990, every state in United States used computers for redistricting, and almost every state used GIS software. The move towards computer-assisted redistricting was also aided by the availability of detailed demographic and geographic data in computer readable formats from the U.S. census Bureau (Maley, 1998)

There are GIS software packages available that are specifically tailored for redistricting. GIS redistricting software provides the following capabilities:

- Draw district plans interactively by assigning geography to districts piece y piece and seeing the results displayed on the computer immediately;
- Produce maps, both on the computer screen and in printed form;
- Generate statistical reports for redistricting plan and each individual district in a plan.

The application of fractal analysis especially in GIS environment provides additional tools to gain insight into the spatial nature of compactness (Knight, 1997). Fractal concepts are one of the techniques that can be applied in geographic study to evaluate spatial characteristics especially in redistricting.

Geographic shapes, such as voting districts, are often complex and have been described by the media, public and the courts as "serpentine", "growing bacteria", "spiders", "shape with tortuous nooks and crannies", and "DNA fragments" (Shiode, 1998; Knight, 1997). All these descriptors are used to relate shape characteristics. The expressive language characterizes geographic shapes as complex in nature. The Random House Dictionary of the English Language (1987) defines a fractal as:

> **"A geometrical or physical structure having an irregular or fragmented shape at all scales of measurement... such that certain mathematical or physical properties of the structure... behave as if the dimensions of the structure are greater than the spatial dimension."**

Fractal geometry is a tool to describe erratic, complex forms of nature which are neither points, lines, or areas but instead fall between the traditional categories of topological dimensions (Muller, 1986). A fractal dimension D, which will fall between one and two (assuming that no geographic polygon is so small that it becomes a point), can describe the complex nature of geographic shape. A fractal dimension allows discussion of values between one, two or three dimensions as summarized in following table.

Table B.3.2: Geographical features with different topological dimensions

Geographical features	Topological Dimension
Point	0
Boundary (Curve)	1
Surface	2
Solid	3

Although different geographical features have different topological dimensions, a complex curve may wander on a surface. In the extreme case, the curve may be so complex that it effectively fills the surface on which it lies. A curve will have a fractal dimension of a real number between one and two. A complex curve that approaches surface filling will have a fractal dimension approaching two. Therefore, the more complex the geographic boundary (curve), the higher the fractal dimension. Boundaries will have a fractal dimension somewhere between a straight line, with a fractal dimension of one, and a boundary that is so complex it is space filling with a fractal dimension of two.

Fractal dimension can be used as a model for measuring the complexity of the district boundary (curve) or district's area (polygon) (DeCola, 1993). The complexity of geographic areas can be described by estimating the fractal dimension (D). The resulting D not only used as a model of complexity but also as proposed by this paper, an alternate measure of compactness.

Moreover, some of the GIS software packages have the ability to customize and automated processing through their scripting languages. Each district was processed based on its unique shape, size, location and whether it contained state boundaries. For each district, the following data was collected:

- Area of the district
- Perimeter of the district
- Size of the largest circles that will fit within the boundary
- Size of the maximum cell side

B.3.6 Fractal Analysis Compares to traditional Compactness Measures

The use of fractal dimension to evaluate the complexity of shape boundaries has the potential of providing the advantages, which are not found in conventional compactness measures. Indeed, fractal dimension appears to be superior in evaluating the full range of boundary complexity with the ability to summarize the district complexity (Shiode, 1998; Knight, 1997).

- Conventional compactness measures use very specific and limited geometric parameters to represent the whole district. This allows specific shape characteristics to unduly influence the resulting measure.

- All conventional measures must consider the geographic shape as a whole. If a geographic area contains a meandering river or oddly shaped state boundaries, it still must be included in conventional measure calculation. Consequently, the measures may be highly influenced by factors that are not included as part of the intended evaluation. Not all fractal measures require the boundary be a closed polygon. This allows the focus of the investigation to be modified. Geographic boundaries that are not subject to study, such as the state boundary or shorelines, may be removed.

Fractal analysis (fractal dimension) may provide a clear approach to analyze the spatial structure and form than current compactness measures. It appears feasible and reasonable to apply fractal analysis to define district complexity or compactness especially with fractal figures which share the three types of features (Shiode, 1998).

B.3.6.1 No characteristic length

Fractal figures are non-differentiable. They are different than geometric shapes, which usually have definite scales and lengths such as radius or circumference of a circle and the edge or diagonal of a square to characterize themselves.

B.3.6.2 Self-similarity

Fractal figures hold self-similarity, their shapes do not change even when observed under different scale. This nature is called scale-invariance. Even the geo-data used in GIS, such as coastline shape and urban space pattern, hold strong statistical self-similarity between different scales. Therefore, when zoom in closer to a landscape image and as the level of detail increases, a pattern similarly will be found at a larger scale level.

B.3.6.3 Non-integer dimension (fractal dimension)

The actual value of these fractal figures differs slightly, depending on the method of defining it. Measure the fractal dimension are as following:

- changing coarse-graining level (box-counting methods),
- using the fractal measure relations,
- using the correlation function,
- using the distribution function, or
- using the power spectrum.

B.3.7.0 Conclusion

In short, redistricting process is important in many areas like political, business, school and so on in order to take control or manipulate the geographical space effectively. Therefore, the measures of geographical shape will be an enormous endeavor will be useful in these areas. However, traditional redistricting measures on spatial features are unsatisfactory to produce the optimal result in terms of their compactness and continuity. Thus, it comes to the idea to use fractal dimension to measure the spatial features in the redistricting application. It is believable that fractal dimension will be an alternative to create an optimum district boundary with its powerful features.

Reference

Agnew, John (1994); *Geography and Regional Science Program: Geographic Approaches to Democratization*; National Science Foundation; Washington.

Altman, Micah (1997a); *What are the Judicially Manageable Standards for Redistricting? Evidence from History*; California Institute of Technology; Pasadena, California.

Altman, Micah (1997b); *Is Automation the Answer? The Computational Complexity of Automated Redistricting*; California Institute of Technology; Pasadena, California.

Altman, Micah (1998); *Redistricting Principles and Democratic Representation*; Ph.D. Thesis; California Institute of Technology; Pasadena, California.

Cain, Bruce (1984); *The Reapportionment Puzzle;* California: University of California Press.

DeCola, L., & N. S. Lam (1993); *Introduction to Fractals in Geography*; Fractals in Geography; Englewood Cliffs, N.J.: Prentice Hall.

Handley, Lisa & Maley, Michael (1997); *Boundary Delimitation Index*;
URL:http://www.aceproject.org/main/english/bd/default.htm

Hearne W. Walter (1999); *Methods to Madness: Alternative Schemes for Single-Member Redistricting*; The Center for Voting and Democracy.

Hoover, Aaron (1997); *Can We Give Children The Education They Need?* ; Wilmington Star-News

Knight , Jeffrey L. (1997); *GIS Based Compactness Measurement Using Fractal Analysis*; George Mason University Fairfax, Virginia.

Liittschwager, John M. (1973); "The Iowa Redistricting System. In Democratic Representation and Apportionment: Quantitative Methods, Measures, and Criteria"; *New York: Annals of the New York Academy of Sciences*; p.p. 221-234.

Maley, Micheal (1998). Boundary Delimitation Index.

Manley, Rachel T. (1995). How Compactness is Measured.

Muller. J. C. (1986); "Fractal Dimension and Inconsistencies in Cartographic Line Representations"; *The Cartographic Journal 23*; p.p.123-130.

National Conference of State Legislatures (1999). Redistricting task Force: State Districting Principles. URL: http://www.ncsl.org/programs/legman/elect/rdrptpres.htm

Newsom J. (1997); *In-Depth: Redistricting Our Schools: Some Facts about School Redistricting*; Depot News; Piedmont Triad, North Carolina.

Ngu Meng Siong (1998); *GIS Based Application for Redistricting the Election Boundaries* (Unpublished); A final year research for Bachelor Degree study of Faculty of Information Technology, Universiti Malaysia Sarawak (UNIMAS).

Shiode, Narushige (1998); *Towards Application Of GIS For Visualizing Fractal Phenomena*; Center For Advanced Spatial Analysis; University College London.

URL: http://www.aceproject.org/main/english/bd/default.htm

URL: http://www.aceproject.org/main/english/bd/gis.htm

URL: http://www.cse.buffalo.edu/~rtmanley/perim.html