CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

The importance of location research is often due to several factors: first, location decisions are frequently made at all levels of human organisation from individuals and households to firms, government agencies and even international agencies. Second, such decisions are often strategic in nature. That is, they involve large sum of capital resources and their economic effects are long term (Daskin, 1995). In the private sector they have a major influence on the ability of a firm to compete in the market place. In the public sector they influence the efficiency by which jurisdictions provided by public services and their ability to attract households and other economic activities. Third, they frequently impose economic externalities such as pollution, congestion and economic development[1]. Fourth, location models are often extremely difficult to solve, at least optimally. Even the most basic models are computationally intractable for large problem instances (Current et al., 2002). Finally, location models are application specific. That is, their structural form (the objectives, constraints and variables) is determined by the particular location problem under study. Consequently, a general location model that is appropriate for all potential or existing applications does not exist (Current et al., 2002).

---

[1] Economic development may be defined as an increase in income of a particular region as a result of newly created economic activity which does not reduce the activity or income of other areas (Hough and Griffin, 1994).

Location allocation models seek the location of facilities and/or services (e.g. schools, hospitals and warehouses) so as to optimise one or several objectives generally related to the efficiency of the system or to the allocation of resources. One of the common ways to classify this type of problem is the dichotomy of the public versus private sector problem. The two problems are different because of the optimisation criteria used in both cases. In private applications, profit maximisation and capture of larger market share from competitors are the main criteria, while social cost minimisation, universality of the service, efficiency and equity are the goals in the public sector. Since these objectives are difficult to measure, they are frequently surrogated by minimisation of the locational and operational cost needed for full coverage of the service, or the search for maximal coverage given an amount of available resources (Daskin, 1995). Here, we are considering the following two models: the first approach that responds to the best locational configuration is a *p*-median problem and the second one is a covering problem (Location Set Covering Problem (LSCP), or Maximal Covering Location Problem (MCLP)). Most public facility location models use one of these approaches (or a combination of both) to set the foundations of the formulation at hand.

Locational analysis is a form of analysis done to investigate where to physically locate a set of facilities so as to optimise some objectives (such as to minimise the cost) of satisfying some sets of demands (customers or users of a service) subject to some sets of constraints. This form of analysis allows a decision-maker to analyse the facility location decisions which will affect a system's flexibility to meet these demands as they evolve over time. Good location decisions are integral to a particular system's ability to satisfy its demand in an efficient manner.

1.2   PROBLEM STATEMENT

In health service planning, one of the tools for location analysis is quantitative location-allocation modelling which plays a significant role as it provides a framework for investigating accessibility problems, comparing the quality (in terms of efficiency) of previous location decisions, and providing alternative solutions to change and improve the existing system (Rahman and Smith, 1999). Proper provision for health is essential for economic development and because of that several studies on analysing and evaluating the implementation of planning for health development have been done in other developing countries (for example, Wang for the World Bank, 2002). As Malaysia strives to be a developed country by the year 2020, the demand for a better healthcare system has become more significant. Currently, the primary healthcare (PHC) service in Malaysia is among the best in the developing countries (Annual Report Ministry of Health Malaysia (MOH), 2005; Hsu, 2005). However, most of the clinics especially in the urban areas are highly congested and the average waiting time for 50 visitors is 102.16 minutes (Utusan online, 12 Jan 2010). The Ministry has also decided to increase the operation hours in order to manage the increasing need for public services. This might have resulted from an improper planning of locating the facilities and/or unexpected growth of demand volume within the area or imbalance growth of demand in a certain area.

The effectiveness of applying location analysis in many developing countries has been studied and yet to the best of our knowledge a comprehensive study of the Malaysian Health System has not been carried out.

1.3   RESEARCH OBJECTIVES AND SCOPE

In the first part of the research, the Malaysian healthcare delivery system and the national policy in determining the efficiency of locating the facilities are studied. There are lot studies of mathematical models that can provide good indication in ensuring that the facilities are located optimally and provide good services to the customers. However, in order to fulfil the target of the healthcare delivery service in the national policy, the study focuses on only two problem models. The problems are chosen based on the objective functions which maximise coverage of the service and also minimise the travelled distance. The reviews on the literature also prove the usefulness of the two problem models in finding an optimum location of the facilities for healthcare services.

The MOH Annual Report of 2005 also highlights that about 30 percent of all the causes of hospitalisation in the hospitals are related to reproduction health such as normal deliveries, complication of pregnancy childbirth and puerpeirum[2] and certain condition originating in the perinatal[3] period. Based on this information, the study is narrowed down to focus only on the location of the facilities which provide the first primary care to the society. The necessary data are collected on both small and moderately large study areas upon the approval from the MOH.

The research objectives are as follows:

1. to investigate the health service accessibility problems, compare the quality (in terms of efficiency) of previous locational decisions, and generate

---

[2] **Puerperal fever** (from the Latin *puer*, *male child (boy)*), also called **childbed fever**, can develop into **puerperal sepsis**, which is a serious form of septicaemia contracted by a woman during or shortly after childbirth, miscarriage or abortion. If untreated, it is life-threatening. From http://en.wikipedia.org/wiki/Puerperal_fever

[3] **Perinatal:** Pertaining to the period immediately before and after birth. The perinatal period is defined in diverse ways. Depending on the definition, it starts at the 20th to 28th week of gestation and ends 1 to 4 weeks after birth. The word "perinatal" is a hybrid of the Greek "peri-" meaning "around or about" and "natal" from the Latin "natus" meaning "born." (At Medicine.net.com, 8 Sept. 2011)

alternatives either to suggest more efficient service systems or to improve the existing systems

2. to examine the effects of constraints on location decisions that are present in real-life decision

3. to propose an alternative solution to constrained location allocation problems specifically capacitated problems (CPMP and CMCLP)

4. to formulate location decision model (by considering the existing national policy for a facility) for public healthcare facilities in Malaysia as well as in other countries

1.4   RESEARCH CONTRIBUTIONS AND REPLICABILITY

The thesis provides an investigation on the health service delivery system and the accessibility problems in the country. It helps in providing a measurement of quality (in terms of efficiency) of the existing public healthcare facilities in locational decisions. Application of the available basic models in determining the optimal locations for facilities based on the national policy is done on the Malaysian data for the first time. The data used for the models are collected primarily from the clinics as no such information is available readily in the healthcare delivery system. From the study, the relevant authority can generate alternatives either to suggest a more efficient service system or improve the existing systems.

More realistic models that introduce real life constraints are also applied into the data of the healthcare delivery system. An analysis of the capacity constraint which imposes on the facility is conducted. This highlights the need to revise the national policy on capacity per facility as the productivity level of service by the individual in

charge as well as the building and equipment also improves. The study also shows the availability of the commercial optimisation software CPLEX in providing a rough idea on the performance of the facility network in serving the population.

The earlier study also showed the limitation of the readily available commercial software; hence an alternative solution is produced. A heuristic method based on Genetic Algorithm (GA) is proposed to overcome this. The algorithm is tested using the available test data and comparing to the existing solution available in the literature. As the algorithm is well known and can easily be adapted to any type of data, this has contributed a new alternative as the solution to the location allocation model.

Aside from the constraint, the real life problem also involves more than one objective function. Single objective models often do not take into account other criteria, which may be equally important in finding an optimal solution to the models. A more suitable model combining more than one objective is introduced in order to take into account the national policy and aim for "Health for All". The model is extended to incorporate time as the facility location decisions are frequently long-term in nature. The models are applied to the Malaysian data and they provide an insight of the potential upgrade of the existing facilities and potential new locations of facilities. The new locations are based on the relevant parameters, such as changes in population structures (increase demand) and an increased ability to treat (increase capacity) in the location decision.

The relevant data collected and analysed are provided in the Appendix for reference. Similarly, the programming code for the algorithm is provided in the Appendix.

1.5   THE THESIS ORGANISATION

The thesis is divided into nine chapters. The field of location analysis, the background of the study and the research objectives are introduced in this chapter.   The contributions of the study are also highlighted in this chapter, together with the information on how the study can be replicated for further research.

Chapter 2 describes the relevant models being used to model the public healthcare facilities and highlights studies conducted by previous researchers. The review includes the suitability of the basic facility location model and the extension to the models together with the solution methods.

Chapter 3 provides an overview of the Health Delivery System in Malaysia. In particular, it gives the description of the nature of the diseases prevalent in the country and the availability of the medical treatment to prevent these diseases. It also highlights the focus of the government's health policy to provide maximum healthcare to all by the year 2010. The health profiles for the selected study areas consisting its location, importance, population volume and number of facilities are detailed in this chapter. The information is used several times throughout the study. The two study areas represent the small network of 179 nodes and the larger network of 809 nodes.

Chapter 4 highlights the initial study of the un-capacitated model. The chapter starts with introducing the previous studies done and the solution methods on the un-capacitated p-median and MCLP. The analysis and set up of the data for the small network in Telok Panglima Garang follows after the literature. Two assumptions are made in conducting the study and a sensitivity analysis on the number of facilities open

within the area is carried out. All analyses are done using the commercial optimisation software, CPLEX 10.2.

In Chapter 5, the limited capacity per facility is added as a constraint to the MCLP and the solution methods to the capacitated MCLP (CMCLP) are discussed. A genetic algorithm based heuristic is proposed to solve the model using the new chromosome representation that combines the number of facilities to open and the order of the node assignment. The performance of the algorithm is benchmarked using data from the literature for the various network sizes. The CMCLP model solved by the GA based heuristic solution is then extended to the small network and the results are compared to the performance of the CPLEX 10.2. As the capacity is limited, a sensitivity analysis on three sets of capacities based on the Malaysian government's national health policy is done. The chapter ends with the extension of the model to the larger network of 809 nodes.

Chapter 6 describes a capacitated p-median (CPMP) model. The chapter starts with discussing the relevant solution methods to CPMP from the literature and introduces the test instances that are used very commonly among the researchers. The GA based heuristic introduced in Chapter 5 has been modified to fit into the CPMP objective function and the performance is compared to other GA based algorithms from the literature. The model is then applied to the small network and the performance is analysed using the three sets of capacities introduced in Chapter 5. As the modified objective function of CPMP in this chapter can also highlight the population coverage volume, the performance is compared to CMCLP.

After the application of the basic models of both the *p*-median and MCLP, Chapter 7 shows the extension of the two models. The bi-objective model that combines the two objectives from the models in Chapter 5 and 6 are solved using the GA based heuristic. The performance of the algorithm is also compared to the Lagrangian solution method and existing GA based algorithm from the literature with different representation. As the two objectives are combined using the weighting method, an analysis of the suitable weight values for the data is carried out. This is followed by the extension to a dynamic conditional CMCLP model. The model locates an existing facility with the potential to be upgraded and/or locates an additional new facility with the pre-specified time such that the percentage of population coverage is maximised. A new formulation is proposed and an analysis on the application to the large network data is done.

Each chapter in the study proposes a model that is suitable to model public healthcare facilities. The suitability of each model to model Malaysian public healthcare delivery planning is discussed in Chapter 8. The strengths and weaknesses of each of the models are compared and highlighted.

Chapter 9 concludes the findings with the contributions of this study and directions for future research.

CHAPTER 2

LITERATURE REVIEW

2.1   INTRODUCTION

Facility Location Modeling is a branch of operations research used to address the problem of locating facilities with regard to the existing facilities and/or clients in order to optimise one or several economics, security, social, environmental and other criteria. It is also concerned with the allocation of demands to facility. In some cases, it is important that demands at a site not to be split between facilities. For example, in some retailing operations, a retail store must be supplied by a single warehouse. For administrative reasons, the store's supply cannot be split among different warehouses. In other cases, such as ambulance service, the demand can be served by any available facility. Facility location models must reflect these different demand allocation policies and must then allocate demands (or fractions of the total demand in a region) to different facilities. Most of the time, the demands will be allocated to the nearest (available) facility; however doing so may not be optimal. This is because each demand area has distinct attributes, such as population density, economic importance, geographical feature, weather pattern and many others. Therefore, different requirement of facility quantity and quality (in terms of distance) should be assigned for each demand point so that all demand points can be serviced in a balanced and optimal manner (Dessouky et al., 2006).

The location allocation model is the method that involves simultaneously selecting a set of locations for facilities and assigning spatially distributed sets of demands to these facilities in order to optimise some specified measurable criterion (Rahman and Smith, 2000). The location problems and models may be classified in a number of ways. The criteria used to classify may be based on the topography that is used, or the number of facilities to be located. Problems may also be classified based on the nature of the inputs (e.g. whether they are static or dynamic, known with certainty or only known in a probabilistic sense). Models may further be classified based on whether single or multiple products or demands must be accommodated by the facilities being located, whether there is one objective or multi objectives, whether the facilities are of unlimited capacity or are capacitated, as well as a variety of other classification criteria.

The classification of the problems is also directly related to its objective. The first and possibly the most important problem in location modeling is to select a suitable objective function (Rushton, 1987). The formulation of an objective depends on the types of organisation whether it is a private or public owned organisation. The private owned organisations often locate their facilities at a location that maximises its profit and/or minimises the cost involved. On the other hand, the most common assumption for public owned organisation that provides service is to locate its facilities at which the social cost is minimised; and/or the social benefit is maximised (Hansen et al., 1980). In the health provision for the public, the objective formulation leads to four basic facility-location models: p-center, p-median, set covering and maximal covering problems. We focus on the choice of the four models mentioned and the review is limited to the studies which have explicitly addressed the problems of a class of public facility location, also popularly known as a central facility location (Hodgart, 1978).

In health provision, the location allocation model is often used to find a set of optimal location patterns for healthcare facilities in a region or to compute an optimal set of new locations to add to an existing set of facilities. They can also be utilised to carry out sensitivity analysis such as to evaluate the effects of constraints on location decisions or to evaluate the effectiveness of past decisions in order to determine the need of future planning. The implications of poor location decisions are increased expenses and/or degraded customer service. If too many facilities are deployed, capital costs are likely to exceed the desirable value. If too few facilities are utilised and/or if they are not located well, this can result in increases in mortality (death) and morbidity (diseases). Even if the correct number of facilities is used, poorly sited facilities will result in unnecessarily poor customer services. Thus, facility location modeling takes on an even greater importance when applied to the sitting of health care facilities (Daskin and Dean, 2004). It can assist the decision makers to evaluate and compare the combinations of the alternative locational goals.

In Malaysia, the objective of locating the health facility is to ensure the full coverage of health accessibility within some allowable distance (Annual Report Ministry of Health Malaysia, 2005, 2006). In general, Malaysia is said to have a nearly universal access to health for most Malaysians, but this is undeclared and not structurally defined (Quek, 72010). The public health sector is also heavily subsidised by the government to the tune of 98 percent of all the total health expenditure (Quek, 2009). Despite being praised by many world authorities that Malaysia's primary healthcare structure is among the best within the developing countries, the clinics are still severely overcrowded, over-utilised and often understaffed (Quek, 2010). Hence, in order to model the Malaysian public healthcare delivery system, the literature has

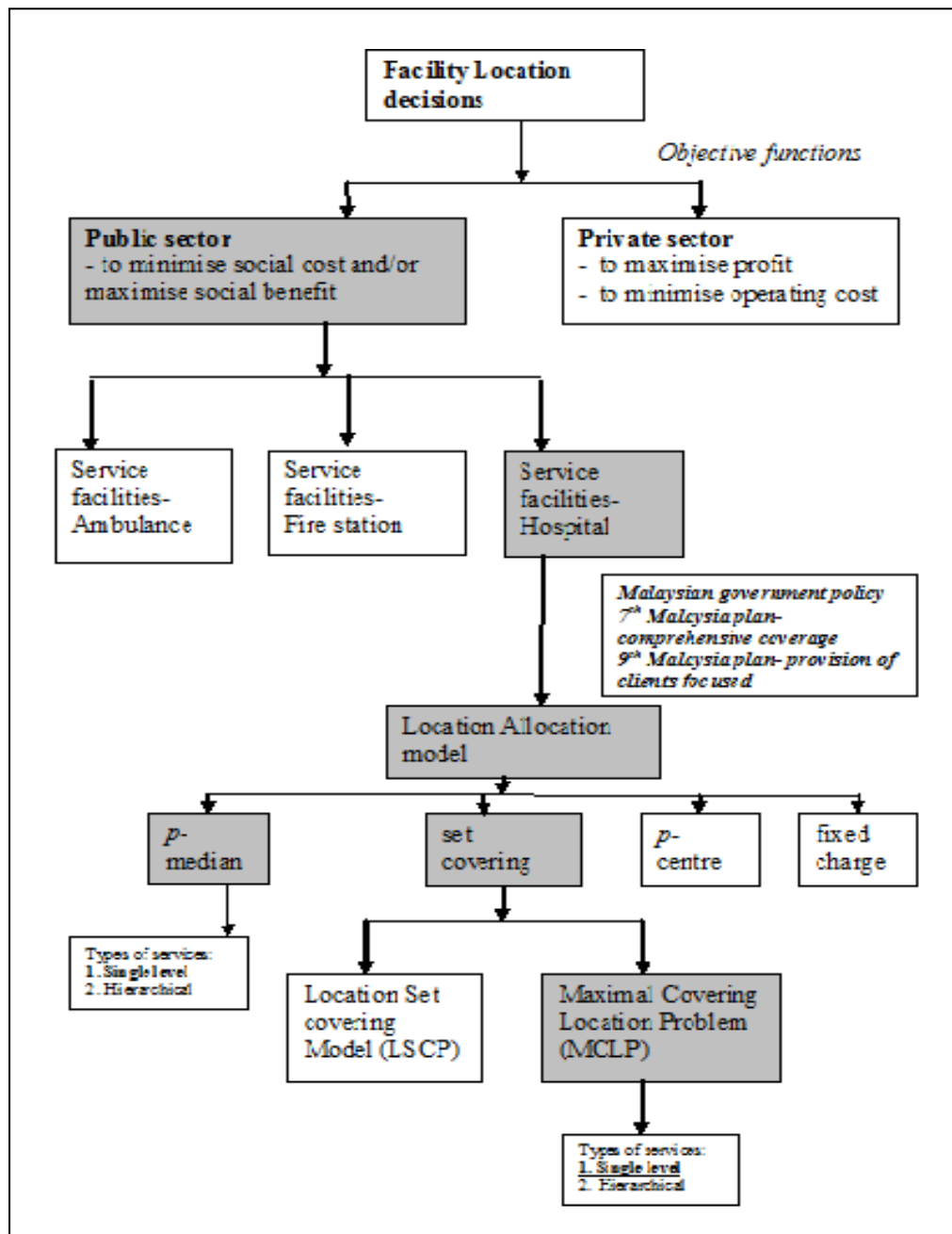been limited to the relevant facility location models as well as extended to relevant model applications.

Figure 2.1: Flow chart of the study

2.2   HEALTH FACILITY LOCATION MODELS

In this section, the basic facility location models that form the basis of almost all models used in healthcare applications is reviewed. All the models are in the class of

discrete facility location models, as opposed to the class of continuous location models. Discrete location models assume that demands can be aggregated to a finite number of discrete points. Thus, a city might be represented by several hundred or even several thousand points or nodes (e.g. census tracts or even census blocks). Similarly, discrete location models assume that there is a finite set of candidate locations or nodes at which facilities can be sited. Continuous location models assume that the demand is distributed continuously across a region. The study is restricted to discrete location models as they have been used far more extensively in healthcare location problems mentioned earlier.

Research for locating health facilities in the context of the problems of developing countries has developed two categories of models. Some research has been directed towards the locations of components of a health care system in which facilities are considered to be of one type (with respect to the level of service provided). These models are referred as single-level location-allocation models. At the same time some researches that consider the level of service provided have been referred as hierarchical location allocation models (Refer to Section 2.2.4).

The following sections introduce the discrete location models that have been used extensively in healthcare location problem. Some typical notations are used in most of the models as follows: The sets I and J represent the clients and sites for facilities respectively. Variable x_{ij} is 1 if client i is assigned to facility j, y_j is 1 if a facility is sited at j or 0 otherwise. a_i is the demand volume at demand node i and d_{ij} is the distance between demand node i and facility j.

2.2.1 *p*-centre model

Given a set of demand nodes, candidate facilities and a fixed number of facilities to be located, the *p*-centre problem attempts to minimise the maximum distance between a demand node and its nearest facility; (Daskin, 1995). Because of this, it is also referred to as the *minimax* problem. The *p*-centre problem seeks the location of the *p* facilities. Each demand point receives its service from the closest facility. This problem is equivalent to covering every point in the area by *p* circles with the smallest possible radius.

The *p*-centre model is formulated as:

Minimise $\qquad Z_1 = W$ $\hfill (2.1)$

Subject to:

$$\sum_{j \in J} y_j = p, \qquad \forall \quad j \in J \hfill (2.2)$$

$$\sum_{j \in J} x_{ij} = 1, \qquad \forall \quad i \in I \hfill (2.3)$$

$$x_{ij} \leq y_j \qquad \forall \quad i \in I, j \in J \hfill (2.4)$$

$$W \geq \sum_{j \in J} d_{ij} x_{ij} \qquad \forall \quad i \in I \hfill (2.5)$$

$$y_j = 0, 1 \qquad \forall \quad j \in J \hfill (2.6)$$

$$x_{ij} \geq 0 \qquad \forall \quad i \in I, j \in J \hfill (2.7)$$

The objective function (2.1) minimizes the maximum distance *W* between a demand node and the closest facility to the node. Constraint (2.2) stipulates that *p* facilities be located. Constraint (2.3) states that all of the demand at node *i* must be assigned to a facility at some node *j* for all nodes *i*. Constraint (2.4) states that demands at node *i* must be assigned to open facility

at *j*. Constraint (2.5) state that the maximum distance *W* between a demand node and the nearest facility must be greater than the distance between any demand node *i* and the facility *j* to which it is assigned. Constraints (2.6) and (2.7) are the integrality and non-negativity constraints, respectively.

The *p*-centre model addresses the problem of needing too many facilities to cover all the demands by relaxing the service standard (i.e. by increasing the coverage distance). This model finds the location of the *p* facilities to minimise the coverage distance subject to a requirement that all demands are covered. In the last several decades, the *p*-centre model and its extensions have been investigated and applied in the context of locating facilities such as hospitals, emergency medical service (EMS) centres, fire stations, and other public facilities. For example, in order to locate a given number of emergency facilities along a road network, Garfinkel et al. (1977) examined the fundamental properties of the *p*- centre problem. He modeled the *p*- centre problem using integer programming and the problem was successfully solved by using a binary search technique and a combination of exact tests and heuristics. ReVelle and Hogan (1989) formulated a *p*- centre problem to locate the facilities so as to minimise the maximum distance within which the EMS is available with some percentage of reliability. System congestion is considered and a derived server busy probability is used to constrain the service reliability level that must be satisfied for all demands.

Stochastic *p*- centre models have also been formulated for the EMS location problems. For example, Hochbaum and Pathria (1998) considered the emergency facility location problem that must minimise the maximum distance

on the network across all time periods. The cost and distance between locations vary in each discrete time period. The authors used k underlying networks to represent the different periods and provide a polynomial-time 3-approximation algorithm to obtain the solution for each problem. Talwar (2002) utilised a p-centre model to locate and dispatch three emergency rescue helicopters to serve the growing EMS demands due to accidents occurring during adventure holidays such as skiing, hiking and climbing the north and south Alpine mountain ranges. One of the model's aims was to minimise the maximum (worst) response times and the author used effective heuristics to solve the problem. There are still many other applications and analyses to various p-centre models. The readers interested in these applications and their mathematical formulations can refer to Handler (1990), Brandeau et al. (1995), Daskin (2000), and Current et al. (2002).

The p- centre model is useful when there are not enough facilities in reality while the service has to cover all the clients within a target region. And it is the most suitable to model the emergency medical system compared to sitting the fixed facilities. Moreover, as the p-centre criterion seeks to minimise the maximum distance, it does not consider the demands of the nodes which receive service from the facility.

2.2.2 p-median model

As mentioned earlier in Section 2.1, the objectives of the public facility locations are more difficult to interpret as they can have a number of possible problem statements. For example, if the problem is to locate emergency

ambulance services, a possible criterion would be to minimise the average distance or time an ambulance must travel in order to reach a random incident or the patient must travel to reach the closest emergency medical service facility. In addition, a more suitable criterion could be to minimise a cost function relating to both travel times and distances and possibly other travel attributes. The criterion that minimises a cost function is referred to a *minisum* criterion and locations that optimise the *minisum* criteria are referred to as medians on networks (Weiss et al., 1971; Rahman, 1991). This model is known as the *p*-median problem.

The *p*-median problem was first introduced by Hakimi (1964) and takes the notion that when the average (total) distance decreases, the accessibility and effectiveness of the facilities increase into account. The model is defined as: Given *p* facilities, determine the locations of these facilities such that the average (total) distance between demands and the locations of the facilities is minimised, Z_2. Later ReVelle and Swain (1970) formulated the *p*-median problem as a linear integer programme and used a branch-and-bound algorithm to solve the problem.

The problem is formulated as follows:

Minimise $$Z_2 = \sum_{i \in I} \sum_{j \in J} a_i d_{ij} x_{ij} \tag{2.8}$$

$$\sum_{j \in J} y_j = p, \qquad \forall \quad j \in J \tag{2.9}$$

$$\sum_{j \in J} x_{ij} = 1, \qquad \forall \quad i \in I \tag{2.10}$$

$$x_{ij} \leq y_j \qquad \forall \quad i \in I, j \in J \tag{2.11}$$

$$y_j = 1 \qquad \text{for all existing facilities} \qquad (2.12)$$

$$x_{ij}, y_j = [0,1] \qquad \forall \quad i \in I, j \in J \qquad (2.13)$$

Constraint (2.9) limits the total number of facilities to equal to p, while constraint (2.10) ensures that all demand nodes are assigned to a facility. Constraint (2.11) guarantees that a demand node is only allocated to an open facility. Constraints (2.12) and (2.13) fix the locations of the facilities that already exist and impose the integrality restriction respectively. Note that the value of p is the total number of facilities, including both existing facilities and facilities that are to be located.

Since its formulation, the p-median model has been enhanced and applied to a wide range of healthcare facility location problems. Among the related studies are: Carbone (1974) formulated a deterministic p-median model to minimise the distance needed to be travelled by customers to medical or day care centres; Berlin et al. (1976) investigated two p-median problems to locate hospitals and ambulances; Carson and Batta (1990) proposed a p-median model to find the dynamic ambulance positioning strategy for a campus emergency services; Mandell (1998) developed a p-median model and used priority dispatching to optimally locate emergency; Paluzzi (2004) discussed and tested a p-median-based heuristic location model for placing emergency service facilities for the city of Carbondale, Illinois, in the United States of America (USA). However, the p-median model does not consider the "worst" case situation in which there are clients who may live very far from this facility. These few remote clients will be forced to travel far and so it may result in inequities.

2.2.3   The Set Covering Model and Maximal Covering Model

When the "worst" case situation mentioned in the previous section is not considered, this may result in the decline of usage of the service facility, when the travel time exceeds some critical value. Therefore, it is reasonable to consider the maximum service distance as a constraint in formulating a location problem. In the facility location, a priori maximum distance is known as the "covering" distance. Demand within the covering distance of its closest facility is considered "covered".   The measurement of the maximum distance is that demand is not fully satisfied even though the facility is closer but it should be within the distance. However, by gradually tightening the maximum distance (time) constraint, a possible outcome is that at some point, the given facilities will become insufficient in number to cover all the points of demands within the distant constraints.

The Location Set Covering Problem (LSCP) model which aims at minimising the number of facilities required to cover all the demand nodes within a specified distance, is proposed in reference to the sitting of the emergency services. This first location covering location problem was introduced by Toregas et al. in 1971. The LSCP is formulated as an integer programming model with the objective to locate the minimum number of facilities required to "cover" all of the demand nodes. The objective is subject to each demand node is covered by at least one facility.

The LSCP is formulated as an integer programming model using the following notations:

$$a_{ij} = \begin{cases} 1, & \text{if site } j \text{ can cover demand } i \\ 0, & \text{if not} \end{cases}$$

$$y_j = \begin{cases} 1, & \text{if site } j \text{ is selected for a facility} \\ 0, & \text{if not} \end{cases}$$

The objective Z_3 is to locate the minimum number of facilities required to "cover" all of the demand nodes. The author's formulation for the LSCP is:

$$\text{Minimise} \qquad Z_3 = \sum_{j \in J} y_j \qquad\qquad\qquad\qquad (2.14)$$

$$\text{Subject to:} \qquad \sum_{j \in J} a_{ij} y_j \geq 1 \qquad \forall \ \ i \in I \qquad\qquad (2.15)$$

$$y_j = \{0,1\} \qquad \forall \ \ j \in J \qquad\qquad (2.16)$$

Constraint (2.15) ensures that each demand node is covered by at least one facility; Constraint (2.16) enforces the yes or no nature of the sitting decision.

As mentioned earlier, some planning scenario exists where there is a desired coverage distance and some maximum distance beyond which service is unacceptable. To solve this problem, the Maximal Covering Location Problem (MCLP) was first introduced by Church and Re Velle (1974). The MCLP seeks to maximize the population which can be served within a stated service distance or time, given a limited number of facilities. In reality, there may not be enough resources to extend coverage to all areas. This leads to a problem whether coverage within a desired service distance can be maximised whilst a fixed number of facilities are to be located. The MCLP can be considered as a budget limited variation of the set covering model. In many facility planning situations, a budget does exist. The model can be used to solve many types of problems in the public service facility planning such as locating fire stations or ambulance

services centres especially in rural areas where the demand is sparse and it is not optimal to cover every household.

The MCLP formulation is having similar constraints as in *p*-median except for the number of facility open to be at least *p* as in Constraint (2.18).

Maximise $\quad Z_4 = \sum_{i \in I} \sum_{j \in J} c_{ij} a_i x_{ij}$ $\qquad\qquad\qquad$ (2.17)

Subject to constraints 2.3, 2.4, 2.4 and 2.6 and;

$$\sum_{j \in J} y_j \le p, \qquad\qquad \forall \quad j \in J \qquad\qquad (2.18)$$

The level of service provided to covered demand is obviously controlled by the variable c_{ij} and it is 1 if the demand volume a_i is assigned to a facility within the coverage distance *S*, where *S* is the maximum service distance or time; however, an uncovered demand node could be assigned to any available facility, regardless of its proximity.

MCLP and its variants have been used to solve many related health facility locations by Eaton et al. (1981) on determining an optimal location of health clinics; Moore and Re Velle (1982) in the planning of health service hierarchies in Honduras; Rahman and Smith (1996) in locating new healthcare facilities in a sub-district, Bangladesh; Verter and Lapierre (2002) in locating preventive health care facilities in Fulton County, Georgia, USA; and most recently, Cocking et al. (2006) in optimally locating new health facilities for better healthcare distribution in the Nouna District, Burkina Faso. The literature is described in details in Chapter 3.

2.2.4 Hierarchical Location Allocation Models

As mentioned earlier, in many developing countries, facilities at different levels offer different types of services. Local clinics may provide basic care as well as diagnostic services. Community hospitals will provide basic care and diagnostics services as well as out-patient surgery and limited in-patient services. Regional hospitals may perform out-patient surgery, in-patient surgery, and provide a full range of in-patient services. However, regional hospitals may or may not provide basic care and diagnostic services. Though, the facilities are distinguished by the services they provided, there exist some sort of link between the facilities being located. In a health care context, there may be linkages that identify which local clinics can refer patients to which community hospitals.

In order to classify the hierarchical facilities, it is often useful to look at the way in which services are offered and by the region to which services are provided by the facilities (Tien et al., 1983). Let the number of levels of possible services in the hierarchy (or the types of facilities) from 1 to m. The first classification can be a successfully inclusive facility hierarchy in which a facility at level m (the highest level) offers services 1 through m. Another classification is a successfully exclusive facility hierarchy in which a facility at level k offers only service type k. A typical hypothetical hierarchical healthcare system is summarized in Table 2.1.

Table 2.1: Hypothetical Hierarchical Health Care System (Daskin, 1995)

Facility	Service Provided			
	Basic Care	Diagnostic Service	Out-patient surgery	In-Patient Surgery
Clinics	Yes	Yes		
Community Hospitals	Yes	Yes	Yes	
Regional Hospital			Yes	Yes

From Table 2.1, note that the community hospitals and clinics constitute a successfully inclusive hierarchy since community hospitals provide all services offered by clinics plus at least one additional class of services, namely out-patient surgery. Regional hospitals and clinics however, illustrate an exclusive hierarchy since the services offered by regional hospitals are not offered by clinics and vice versa. Note, however, that it is not illustrative of a successively exclusive hierarchy since the two levels are separated by another level, that of the community hospital. Finally, the relationship between regional hospitals and community hospitals is more complicated and is neither successively inclusive nor successively exclusive.

Most hierarchical facility location problems have been posted as variants of p-median models. Using the following notations, as an example, a model for a successively inclusive facility hierarchy can be formulated.

Inputs

h_{ik} is demand for type k services at node i

d_{ij} is distance between node i and candidate location j

p_k is number of type k facilities to locate

Decision variables

$$y_{jk} = \begin{cases} 1 \text{ if a facility of type } k \text{ is located at candidate site } j \\ 0 \text{ if not} \end{cases}$$

$$x_{ijk} = \begin{cases} 1 \text{ if demands at node } i \text{ for type } k \text{ services are satisfied by a facility} \\ \quad \text{at candidate site } j \\ 0 \text{ if not} \end{cases}$$

$$\text{Minimize } Z_5 = \sum_i \sum_j \sum_k h_{ik} d_{ij} x_{ijk} \tag{2.19}$$

Subject to:

$$\sum_j x_{ijk} = 1 \qquad \forall i, k \tag{2.20}$$

$$\sum_j y_{jk} = p_k \qquad \forall k \tag{2.21}$$

$$x_{ijk} \leq \sum_{h=k}^{m} y_{jh} \qquad \forall i, j, k \tag{2.22}$$

$$y_{jk} = 0, 1 \qquad \forall j, k \tag{2.23}$$

$$x_{ijk} = 0, 1 \qquad \forall i, j, k \tag{2.24}$$

The objective function (2.19) minimizes the demand-weighted total distance. Constraint (2.20) stipulates that all demand types at all locations must be assigned to some facility. Constraint (2.21) limits the total number of types k facilities located to p_k. Constraint (2.22) is the linkage constraint that means demands for type k service that originates at node i cannot be assigned to a facility at node j unless there is a type k or higher-level facility located at node j. Note that h is the maximum number of facility types under consideration. Constraints (2.23 and 2.24) are the integrality constraints. Note that the integrality constraint (2.24) associated with the allocation of variables x_{ijk} may be relaxed since demand at node i for service of type k will naturally be

assigned to the closest facility which can serve such demands. In the formulation, Constraint (2.22) links the location y_{jk} and the allocation x_{ijk} variables. They also link the different levels of facilities together. Technically this formulation allows multiple facility types to be located at the same candidate location.

With the following additional notations, a hierarchical maximum covering location problem can be formulated as follows (Daskin, 1995):

Inputs

$$a_{ij}^{kq} = \begin{cases} 1 & \text{if a type } q \text{ facility located at candidate site } j \\ 0 & \text{if not} \end{cases}$$

Decision Variables

$$y_{jq} = \begin{cases} 1 & \text{if we locate a type } q \text{ facility at candidate site } j \\ 0 & \text{if not} \end{cases}$$

$$z_{ik} = \begin{cases} 1 & \text{if demand for service } k \text{ at node } i \text{ are covered} \\ 0 & \text{if not} \end{cases}$$

Maximize $\quad Z_6 = \sum_i \sum_k h_{ik} z_{ik}$ \hfill (2.25)

Subject to

$$\sum_j x_{jq} = p_q \qquad \forall q \tag{2.26}$$

$$z_{ik} \le \sum_j \sum_q a_{ij}^{kq} y_{jq} \qquad \forall i, k \tag{2.27}$$

$$0 \le z_{ik} \le 1 \qquad \forall i, k \tag{2.28}$$

$$y_{jq} = 0, 1 \qquad \forall j.q \tag{2.29}$$

The objective function (2.25) maximizes the total number of demand of all types that are covered. Constraint (2.26) stipulates that exactly p_q types of q facilities are to be located. Constraint (2.27) ensures that demands for service k at node i cannot be counted as being covered unless we locate at least one facility at one or more of the candidate locations which are able to provide service k to demand node i. Constraints (2.28 and 2.29) are the non-negativity and integrality constraints, respectively. Note that the coverage variables z_{ik} need not be explicitly constrained to take on only integer values. In some situations, demands at each demand node can only be counted as being covered if all the services are demanded at the node are covered and can be formulated by having the additional constraint to consider the situation.

In this study, however, we will focus only on the first point of contact for primary healthcare that provides the basic care to the population, which applies to a single level location allocation model. Primary care is the basic or general healthcare that focuses on the point at which ideally a patient first seeks assistance from the medical care system. It also forms the basis for referrals to a higher level of care.

2.3   EXTENDED LOCATION MODELS

All the models mentioned in the earlier sections are basic location models to which most applications are based on. The healthcare location literature tended to address three major topics: accessibility, adaptability and availability (Daskin and Dean, 2004). Accessibility is defined as the ability of patients or clients to reach their healthcare facility; availability focuses on short term balance between the ever

changing demand for services and the supply of those services; and adaptability means to consider multiple future conditions and try to find good compromise solutions. The need of accommodating changes in the real world has directed the facility location research to extend the basic models to more applicable forms. For example, in the earlier sections, the assumption for all models is each facility will be able to serve as many clients as possible. This assumption is valid in many location allocation planning settings; however, there certainly exists situations in which this assumption severely limits the system's ability to provide an effective service. This leads to a revised model that includes the capacity of the facility as a constraint. For example, Jacobs et al. (1996) used a capacitated p-median model to optimise collection, testing and distribution of blood products in Virginia and North Carolina; Cocking et al. (2006) used a capacitated covering model to locate health facilities in Burkina Faso; and Jia et al. (2007) used the capacitated MCLP to formulate large scale emergency facility location problem.

There is also a need to simultaneously locate a number of different services like local clinics, community health centres and regional hospitals. In this scenario, a hierarchical location model is applied. The healthcare delivery system that consists of several levels of service is called hierarchical. In this system, medical centres provide specialised care in addition to the services available at hospitals and clinics. Likewise, hospitals provide more services than are available in health and rural clinics. Here, rural clinics are assumed to be the first point of contact between the rural population and the health system. A patient who is not cured at a clinic may be referred to a hospital; and a patient who goes to a hospital may be referred to a medical centre. The hierarchical model is used more often in modeling the allocation of basic life support vehicles or emergency services. For example, Calvo and Marks (1973) constructed a p-median

model to locate multi-level health care facilities that included central hospitals, community hospitals and local reception centres. The model sought to simultaneously minimise distance and user costs, and maximise demand and utilisation. Later, the hierarchical *p*-median model was improved by Tien et al. (1983) and Mirchandani (1987) by introducing new features where two indices were introduced to distinguish between demand types and facility types, and allowing various allocation schemes to overcome the deficient organisational problem across hierarchies.

A model that attributes to a single objective often does not capture many important elements that are vital in designing effective locational decisions. Many interests of the decision makers prompted the researchers to develop a multi objective model that considers more than one factor simultaneously (Cohon, 1978; Goodchild and Noronha, 1987). Some of the work that have incorporated multi-objective include those such as Haghani (1996) who presented a mixed-integer linear programming model with two objectives; Jayaraman (1999) who developed a multi objective logistics model for a capacitated service facility problem that minimised three objectives: (1) the fixed investment cost, (2) variable operating cost and (3) service attribute in terms of average response distance (or time); Doerner et al. (2007) who presented a multi-objective combinatorial optimisation formulation for a location-routing problem in mobile healthcare management; Chan et al. (2008) who located up to *p* signal receiving stations for demands (or actually distress signals and /or transmission from any target) and was modeled as a multi-objective linear integer programme (MOLIP); Hosseini and Ameli (2011) who formulated a bi-objective model for an emergency service location allocation problem with maximum distance constraint. These studies will be further discussed in Chapter 7.

Facility location decisions are frequently long-term in nature and involve large capital outlays. Hence, there may be considerable uncertainties regarding the way in which the relevant parameters in the location decision will change over time. Uncertainties have also been considered in many location models. Mirchandani (1980) examined a *p*-median problem to locate fire-fighting emergency units with consideration of the stochastic travel characteristics and demand patterns. The author took into account the situations that a facility may not be available to serve a demand and used a Markov process to create a system in which the states were specified according to demand distribution, service and travel time, and server availability. Serra and Marianov (1999) implemented a *p*-median model and introduced the concept of regret and minimax objectives when locating a fire station for emergency services in Barcelona. Location models that addressed the uncertainty problems in which facilities are opened and/or closed over the planning horizon were clearly dynamic. Such models typically resulted in a schedule or plan for opening and/or closing facilities at specific times and locations in response to changes in parameters over time.

These location models have also been extended to solve emergency service location problems in a queuing theory context. An example is the use of the *p*-median model in the Stochastic Queue Median (SQM) as in Berman et al. (1985). Many of the stochastic studies have concentrated on the *p*-median compared to the MCLP. The basis for covering models is to assume that a point is covered if it is within a certain distance from a facility and not covered otherwise. However, as the cover gradually declines, the stochastic gradual cover model is more suitable. Drezner et al. (2010) develop a stochastic gradual cover model that assumed the short and long distances employed are random variables. This refinement of the gradual cover models provides yet a more realistic depiction of the actual behaviour in many situations.

2.4 LOCATION ALLOCATION MODEL FOR PUBLIC HEALTHCARE FACILITIES

Most public facility located models use either the *p*-median problem or MCLP or a combination of both, to set the formulations at hand. This is because these two approaches are adequate for its use in applications in the public sector, as it tends to generate certain equity in the access to the facility by their users (Murray and Gerard, 1997; Smith et al., 2008). Equity in health (health status) means the attainment by all citizens. Equity in health care means that health care resources are allocated according to need; health care is provided in response to legitimate expectations of the people; health services are received according to need regardless of the prevailing social attributes, and payment for health services is made according to the ability to pay (WHO, 2000). Factors from the supply side that influence equal access in healthcare are generally that health care resources must be distributed to regions according to population size, local input (e.g. labour and capital) costs, healthcare needs and (if income affects access) the income mix within each regional population. Additionally, efforts ought to be made to overcome any "inequitable" capacity constraints in disadvantaged areas, to ensure that there are incentives/directives for sufficient facilities and staff to locate and remain within these areas (Oliver and Mossialos, 2005).

This study employs the *p*-median and MCLP as the basic problems to model the location of the public healthcare in Malaysia. These two approaches have several advantages. The first model, the *p*-median, is attractive since the smaller the total weighted (which is equivalent to the average) traveled distance (or time), the more convenient for one to get to the nearest facility. On the other hand, the MCLP determines what can be covered by a given number of facilities and it can also be easily

used to analyse the marginal coverage associated with adding one or more new facilities.

There are several studies conducted in applying the location allocation model for healthcare facilities in developing countries. Gould and Leinbach (1966) studied the problem of locating hospitals and determining their capacities (in terms of number of beds) in the western part of Guatemala, with an objective to minimise the distance traveled between eighteen population centres and three-to-be located regional hospitals. To find the site and size of the regional hospitals based on the existing road networks, the authors considered a *p*-median problem solved using a transportation algorithm. In addition, Mehretu et al. (1983) conducted a study to locate rural health clinics in the Eastern Region of Upper Volta, (now Burkina Faso) with an objective to minimise average distance traveled subjected to the constraint that no one travels more than 5 km. The problem was defined as a *p*-median problem with maximum distance constraint. A modified *p*-median model which addressed accessibility and physician availability at the clinics simultaneously was used to solve the location problem in Mafraq, a district in Jordan (Tien and El Tell, 1984). This study demonstrated the need for improvement in the allocation of the villages in the district to the clinics and the allocation of clinics to the existing health centres.

A similar study by Rahman and Smith (1996) in the deployment of health facilities in rural Bangladesh found the optimal locations of facilities for Health and Family Welfare Centres (HFWCs) in Thangail Thana, Bangladesh. The study resulted in reducing the average distance traveled between villages and facilities by at least 26 percent. This study actually measured the effectiveness of past locational decisions. The percentage of coverage was emphasised in a study in Colombia by Eaton et al.

(1981), in which the MCLP model was used to identify new sites to add into the existing system. Demonstrating the use of location analysis in healthcare planning in rural Colombia, Bennett et al. (1982) utilized the MCLP to determine the number of rural health centres from which (and for which) personnels were recruited as health workers. These centres also served as ambulance bases. The findings suggested only 24 health centres were necessary to have 90 percent of the population covered. Rahman and Smith (1999) also formulated his study in Thangail Thana, Bangladesh as MCLP, in order to find the optimal sites for locating the new facilities to be added to the existing health provision system. This study also suggested that the implementation of a solution of a geographically unconstrained problem in which all facilities are located simultaneously by solving one global problem would be more efficient than the solution suggested by geographically constraint problem (when the facilities are located by taking one service area at a time). The study suggested that the implementation of one of the solutions of the unconstrained problem would make the health delivery system 60 percent less costly, while serving the entire population with a maximum travel distance of 2 km.

Some studies, for example, have combined both the *p*-median and MCLP to solve their problems. A study on the effect of changes to communication links due to the rainy season in a tropical country, specifically concerning the location of health facilities in the Suhum district, Ghana, was carried out by Oppong (1996). The problem was solved as both the *p*-median as well as the MCLP opened the ways to multi criteria decision analysis. Cocking et al. (2006) in their study of optimally locating new health facilities in the Nouna district, Burkina Faso, used the covering model that solved the problem of time and effort required in traveling to a health facility (especially during the rainy season).

The location allocation study was also done in developed countries. Bennet (1981) used the location allocation model to determine facility locations and associated user allocations for primary care health centres established in the Lansing, Michigan area in the USA. Based on the distribution of the un-doctored household members reported in a mail-out survey, a heuristic location-allocation algorithm was used. Beyond simply identifying desirable health care centre locations, the analyses showed that four facilities, rather than the originally proposed five, would yield more tenable and equitable utilisation levels. The allocation results also indicated a preferable sequence for facility development based on differences in the expected utilisation. Another relevant study was carried out by Yeh and Hong (1996) on the use of location allocation model for public facility planning. The model was used to identify the optimal sites for open space and to evaluate the location of the existing open space in Hong Kong. In a developing country, Smith et al. (2008) presented a study of the location modeling for community healthcare facilities in Leeds, United Kingdom. The hierarchical location models based on both the *p*-median and maximal covering were used to assess different possibilities of patients' access in a community situation. The study was hoped to give impartial credence to decisions taken by the authority in locating the new facilities.

2.4.1 The solution approaches

The solution approaches can be categorised into two types: the exact solution approach and the heuristic approach. As the facility location problem is *NP*-hard the solution to a realistic sized problems consumed a large amount of computational resources, resulting in more heuristic based solutions to be proposed and used. The exact solution approach, such as branch and bound, can

produce the best solution but cannot handle models with large amounts of constraints and variables (Gu et al., 2010).

The *p*-median and its extensions are used to model real world situations more often compared to the MCLP. Hence, there are more solution approaches being developed for the *p*-median, such as the exact methods by Beasley (1985), Hanjoul and Peeters (1985), and Brandeau and Chiu (1989). There are also some classical heuristics for the *p*-median: Greedy (Kuehn and Hamburger, 1963), Alternate (Maranzana, 1964) and Interchange (Teitz and Bart, 1968). The three heuristic methods have been extended to several hybrids that improve the solution. Similarly, the algorithms are also used to solve the related location problems such as covering problems. For example, a greedy adding algorithm is studied and used extensively to solve the *p*-median as well as set covering problems: Church and Re Velle (1974), Chvatal (1979), Hodgson (1990), and Daskin (1995). Some modern heuristics are also used to solve the *p*-median and set covering problem like: simulated annealing (Friesz et al., 1992; Chiyoshi and Galvao, 2000); tabu search (Glover and Laguna, 1993); multi-stage algorithm (Rosing et al., 1999); variable neighbourhood search (Hansen and Mladenovic, 1997) and genetic algorithm (Alp et al., 2003). Details of the solution approaches will be given in the following chapters.

2.5   SUMMARY AND CONCLUSION

The *p*-median and MCLP have been identified as the most suitable approaches in modeling the public healthcare facilities in Malaysia as the two approaches have several advantages. The models are also adapted to meet the Malaysia National Health Policy objectives which will be described in Chapter 3. Figure 2.1 summarises the flow of the study based on the basic classification of facility location problem.

Both the *p*-median and MCLP are known to be *NP*-Hard and this has resulted in the development of heuristic solution techniques in an effort to solve large-scale problems to near-optimality with a reasonable computational effort. These methods will be discussed in detail in Chapter 4.