CHAPTER I

INTRODUCTION

Operations research (OR) tries to help decision makers, within organizations, to make better choices. It uses advanced quantitative techniques and is not only useful in production industry but also in a wide array of service provider activities. Today with the continuous improvement of computational tools more difficult problems can be solved and OR has found newer and even more complex applications which range from organizing flight paths to identifying patterns in data mining. (Walsh, 2014) Facility location problems have been studied using different operations research methods. The importance of this research is due to the ample area of application of these problems. In particular, this work proposes a method to find primal and dual bounds for the capacitated p-median problem, as a facility location problem it has many real world applications in areas such as telecommunications, industrial transportation, distribution networks (Santos Correa, Steiner, Freitas, & Carnieri, 2004) (Shamsipoor, Sandidzadeh, & Yaghini, 2012), political districting and sales force territories design (Kaveh, Zadeh, & Sahraeian, 2010), clustering (Yaghini, Momeni, Sarmadi, & Ahadi, 2013), among others. Because of this, location problems have received much attention from the research community (Melo, Nickel, & Saldanha-da-Gama, 2009). In general we could explain a facility location problem as one that is aimed to physically locate a number of facilities in such a way that the total costs of satisfying the customer demands are minimized. In other words, it is desired to provide services to a spatially distributed set of demand points in the best possible way (Senne & Lorena, 2003); this objective must be achieved while subjected to certain constraints. It is important to mention that cost could not only be money but also time, distance, etc. The p-median problem belongs to the facility location family of problems. (Hale & Moberg, 2003).

Within the facility location problems there are two objective functions that are predominantly studied: those that minimize the sum of the total distances, which are known as median problems; and those where the objective is to minimize the maximum distance, known as center problems. While these problems are related, that difference between the p-median and the p-center problem create the need of different approaches when finding a solution to them. We also take in consideration that, in order to find a solution, the problems can be represented in three ways: continuous, discrete and in the form of networks (Hale & Moberg, 2003). It is also important to mention that as noted by Hale & Moberg, "The Hakimi property states that for the *p*-median problem on a network (and therefore for any other median problem), at least one of the alternative optimal solutions will consist entirely of vertices of the network". A center problem is less restrictive by letting the optimum not to only be in vertices but also in edges although this increases the size of the feasible area where solutions can be found. (Hakimi, 1964)

Some authors consider the p-median problem, and other related problems, to be a rather simplistic way of representing real life situations. For example, the *p*-median problem makes use of deterministic parameters such as demands and costs, when in reality both of them tend to vary over time within the same year; the model is valid for a single-period planning horizon, for a single product, one type of facility and location-allocation decisions; while in real life these patterns do not stand and could be seen as insufficient to explain the processes that occur in the real world. Even more, the basic *p*-median problem assumes that all the possible setup sites have the same fixed costs, when this is not the case, the objective function is altered in order to reflect these differences in setup costs. (Melo, Nickel, & Saldanha-da-Gama, 2009)

It has already been mentioned that some problems are related and contain similar characteristics, capitalizing on this some problems will be presented in order to make the capacitated p-median problem more comprehensible in its nature.

In the 1-median problem the goal is to locate a facility in such a place that the sum of distances from the demand points to the facility is minimized. To provide an example for this problem, we could imagine a small community that needs a new marketplace as the current one is too old and small; the current one will close once the new one starts working. The community has not grown homogenously in the last years so the last marketplace location is no longer in the geographical center of the town, so a new place needs to be selected. It is desired that the new marketplace is in the most geographically central part of the town as possible. It has to be noted that depending on the nature of the problem the cost function (the one that is desired being minimized) could be calculated differently not considering just distances but other characteristics. (Elgindy & Keil, 1992)

The p-median problem is related to the last one, but instead of opening just one facility, it is desired to open p (a given number) facilities. There is a set of potential locations where the facilities can be placed. There is also a cost related to assigning a demand point to a facility and the goal remains the same one: to minimize the sum of costs of serving the demand points. The cost of opening facilities can also be considered in the cost function. (Byrka, Fleszar, Rybicki, & Spoerhase, 2015) (Li, 2014) This problem has been widely studied and is a classic in combinatorial optimization. As stated by Senne & Lorena, 2003 "The objective is to locate *p* facilities such as the sum of the distances from each demand vertex to its nearest facility is minimized".

It is also important to note that the p-median problem is NP-hard, even when its network is simple in structure. (Kariv & Hakimi, 1979) Because of that, finding the optimum by exploring all the possible solutions is an unfeasible task, which is the reason why a number of solution algorithms have been developed during years of research.

Mathematical modeling is a representation of real world processes and phenomena in mathematical terms and while they are simplified (Dym, 2004) sometimes it is not necessary to oversimplify a model. The p-median problem assumes that the facilities are uncapacitated, this means that there is no limit in the amount of demand points that a single facility can provide service to. But most of times this is an unrealistic assumption; this is why a capacity is considered for each facility and a measure of demand is assigned for each demand point. (Elgindy & Keil, 1992)

In the literature the capacitated p-median problem is referred as CPMP, also as capacitated warehouse location problem and sum-of stars clustering problem, among others. Regardless of its name, this problem consists in partitioning a set of *n e*ntities into *p*-clusters; the goal is to minimize the distances within each cluster while meeting the constraint of maximum capacity in each cluster. (Maniezzo, Mingozzi, & Baldacci, 1998)

The purpose of this work is to present a method for finding primal and dual bounds for the capacitated p-median problem. A Lagrangean method is used to provide lower bounds. Two different approaches are used to provide upper bounds. All methods are tested with a set of instances found in the literature. The computational results show that the second method proposed to find upper bounds accelerates the convergence of the upper bound and the results show that when it is used better upper bounds are found for most cases or at least it founds the same upper bound in small instances, but it never provides worse upper bounds.

The rest of this document is organized as follows. Chapter 2 is a literature review of diverse works that use different approaches to solve the capacitated p-median problem. Chapter 3 is a theoretical framework which will introduce the concepts necessary to understand the solution method presented in this work. After this, Chapter 4 explains the methodology used in this work to solve the capacitated p-median problem. Then, Chapter 5 shows the computational results obtained. Finally, the conclusions and future research are discussed in Chapter 6.