Discrete Applied Mathematics 164 (2014) 68–85

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

The location-dispatching problem: Polyhedral results and content delivery network design

Philippe Chrétienne [a], Pierre Fouilhoux [a,*], Eric Gourdin [b], Jean-Mathieu Segura [a,b]

[a] Laboratoire LIP6, Université Pierre et Marie Curie, 4 place Jussieu 75005 Paris, France

[b] Orange Labs, France Telecom R&D, 38-40 bd du Général Leclerc, 92794 Issy-les-Moulineaux Cedex 9, France

ARTICLE INFO

Article history:
Received 15 October 2010
Received in revised form 30 March 2012
Accepted 17 May 2012
Available online 7 June 2012

Keywords:
Location
Integer program
Polytope
Branch-and-cut algorithm
CDN design

ABSTRACT

Let $G = (V, A)$ be a directed graph and F be a set of items. The Location-Dispatching Problem consists of determining subsets $L_i \subseteq F$ located at nodes $i \in V$, minimizing the sum of two costs: a piecewise linear installation cost associated with L_i and an access cost for each node of V to reach a copy of each item of F. We formulate this problem as a linear program with binary variables x and integer variables z. We propose a facial study of the associated polytope and we introduce the so-called integrity hop cost inequalities that force z to be an integer as soon as x is binary. Using this, we devise a branch-and-cut algorithm and report some experimental results. This algorithm has been used to solve Content Delivery Network instances in order to optimize a Video On Demand (VoD) system.

1. Introduction

The precise study of combinatorial location problems allows to answer many questions arising when one wishes to determine how many equipments to deploy, where to install these equipments and how many items must be dispatched in each of these equipments, in order to satisfy the client requests. Such a location model is, in particular, well designed to analyze situations where the optimal configuration is a compromise between high installation costs (when equipments are located everywhere) and high access costs (when very few equipments are installed and a client is far away from requested items). In this article, we will study a particular location problem to gain some useful insight in the decision process required to deploy a set of storage equipments in a telecommunication network.

Let $G = (V, A)$ be a directed graph of n nodes and F be a set of m items. Nodes of graph G represent both the clients and the potential places where copies of each items have to be dispatched.

In this paper, an equipment will be able to store at most μ items, with $1 \le \mu \le m$. Such an equipment will be denoted by a μ-batch, that is to say a subset of at most μ items. We will associate with every node i of V a subset of (distinct) items $L_i \subseteq F$. In order to store subset L_i at node i, L_i must be partitioned into a minimum number z_i of μ-batches. Consequently, z_i is required to be an integral multiple of μ. By setting $C_i \in \mathbb{R}_+$ the *installation cost* of a μ-batch at node i, the total installation cost of a subset L_i is then the piecewise linear function $C(L_i) = C_i z_i$, with $z_i = \left\lceil \frac{|L_i|}{\mu} \right\rceil$.

Moreover, the total number of μ-batches, $\sum_{i \in V} z_i$, must be at most p where p is a given integer such that $p\mu \ge m$.

* Corresponding author. Tel.: +33 1 44 27 87 96; fax: +33 1 44 27 70 00.
E-mail addresses: philippe.chretienne@lip6.fr (P. Chrétienne), pierre.fouilhoux@lip6.fr (P. Fouilhoux), eric.gourdin@orange-ftgroup.com (E. Gourdin), jean-mathieu.segura@lip6.fr (J.-M. Segura).

0166-218X/$ – see front matter
doi:10.1016/j.dam.2012.05.001

We denote by $c_{ijf} \in \mathbb{R}_+, i \in V, j \in V, f \in F$, the *access cost* for a node j to reach an item f dispatched in the subset L_i of a node i. Let $d_{i,j} \geq 0$, $(i, j) \in A$, be the distance between nodes i and j in the layout graph G. For a given item $f \in F$, the access cost c_{ijf} is proportional to the minimum length from j to i in G and then $c_{iif} = 0$.

The *Location-Dispatching Problem* (LDP) consists of determining subsets of items $L_i \subseteq F, i \in V$, and assigning a node $i \in V$ to each pair $(j, f), j \in V, f \in F$, such that $f \in L_i$ and $\sum_{i \in V} z_i \leq p$, so that the sum of installation and access costs is minimum, i.e.

$$\sum_{i \in V} C_i z_i + \sum_{j \in V} \sum_{f \in F} \min\{c_{ijf} \mid i \in V \text{ and } f \in L_i\}$$

is minimum.

We will call a node $i \in V$ *median* if $L_i \neq \emptyset$. In fact, solving an LDP instance will correspond to give to each client j a way to reach a copy of each item among the μ-batches located on median nodes. Notice that a solution must satisfy $\cup_{i \in V} L_i = F$ but the subsets L_i need not to be a partition of F.

Many direct applications of the Location-Dispatching Problem can be found in industry when location, production allocation and client assignment decisions have to be taken simultaneously. This is the case, for instance, when a company needs both to locate new equipments (median nodes) and dispatch the products (items) that should be produced or stored inside each of them, with respect to the cost of the product access for their clients. In this case, in the layout graph $G = (V, A)$, the nodes of V correspond both to the clients and the potential facility locations, and the arcs of A represent the possible roads between them. Note that if a node cannot support a facility, the cost C_i can be set to a very high value.

In this article, we will focus on an application of the Location-Dispatching Problem to *Content Delivery Network* (CDN) design [4]. A Content Delivery Network is a system where a particular content (files, videos, etc...) is distributed over several interconnected computers and servers instead of being stored in a single central server (in a classical client–server architecture). The objective of such a decentralized structure is to reduce the load on backbone links, possibly avoiding congestion in bottleneck links around the central server, and, as a result, improve the QoE (Quality of Experience) for the clients. The challenge of CDNs is to distribute efficiently the content as closely as possible to the clients so that their future requests will be fulfilled by these closely located servers. In this case, the layout graph $G = (V, A)$ corresponds to the physical telecommunication network where the nodes of V are both clients and potential server locations, and of the arcs of A are the telecommunications links between them. In particular, a CDN can be used to build a Video on Demand (VoD) system where the videos are replicated and stored on several servers. In fact, an LDP solution corresponds to the decision to locate servers containing copies of the movies of F so that each client has a close access to a copy of each movies. A server on a node can be equipped with several hard-disks, each one of capacity μ. The resulting cost is the sum of the installation cost of the servers (μ-batches) and the access cost for each client to reach each movie.

The LDP can be seen as a particular facility location problem. In the well-known *Uncapacitated Facility Location Problem* (UFLP) (also called Simple Plant Location Problem), some facility locations are to be chosen among a set of candidates and each customer must be allocated to a facility, in such a way that the total (installation plus allocation) cost is minimized. When $m = 1$, LDP is exactly UFLP and is then NP-hard [8]. The survey [11] gives integer formulations for several facility location problems when several types of products have to be managed. In this survey, LDP is not introduced but can be seen as a particular case of the so-called *Multi-Product Uncapacitated Facility Location Problem*. To the best of our knowledge, LDP was not treated before in the literature.

When the number of facilities is fixed to p, UFLP is then called the *p-median problem* (PMP) which is also NP-hard [10]. In fact, when $m = 1, C_i = 0$ and $c_{iif} = 0$, for $i = 1, \ldots, n$ and $f = 1, \ldots, m$, LDP is exactly the p-median problem. A complete polyhedral and experimental study on the p-median polytope can be found in [5,3].

In the first section we will present the model and introduce notation, then, in Section 2, we will devise a linear formulation with binary variables x and integer variables z. Our polyhedral results are given in Section 3 together with new facet defining inequalities, called *integrity hop cost* inequalities, that force z to be integer as soon as x is binary. In Section 4, we will focus on a branch-and-cut algorithm and give some experimental results. In Section 5, we will show how the Location-Dispatching problem may be used for solving a CDN design problem.

2. Mathematical model

Let (G, F, p, μ, c, C) be an LDP instance where $G = (V, A)$ is a graph with n nodes, F is a set of m items, μ and p are integers such that $1 \leq \mu \leq m$ and $p\mu \geq m$, $C_i \in \mathbb{R}_+$ is the cost of locating a μ-batch of items at node i and $c_{ijf} \in \mathbb{R}_+$ is the access cost for a node j to reach an item f dispatched at node i.

We first give a graphical representation of the LDP instance (G, F, p, μ, c, C). Let K_n^m the complete directed multi-graph with n nodes and, for every pair $(i, j) \in V \times V$, m parallel arcs. Each arc is labeled with a distinct item f of F. In K_n^m, an arc $(i, j, f), i \in V, j \in V, f \in F$, will correspond to a median node i which gives to node j an access to item f. A graphical representation of an LDP solution can then be represented by a subgraph $H(V, D)$ of K_n^m where, for each node i and each item f, there is an unique arc (i, j, f) that indicates the minimum cost access given by a median node i to node j for the item f. A loop arc (i, i, f) in subgraph H then corresponds to the fact that node i has a direct access to item f in its own L_i subset.

Fig. 1 gives an example of such multi-graph which corresponds to an LDP solution on 5 nodes and 3 items with $\mu = 2$. In this solution, there are 3 median nodes indicated by squares and 2 non-median nodes indicated by circles. The solid (resp.