Robust Multi-source multi-commodity capacitated Facility Location Problem (cFLP)

Dimitri Papadimitriou, Piet Demeester
Email: dimitri.papadimitriou@alcatel-lucent.com, p.demeester@ugent.be

Abstract

Given a set of potential facilities F (with individual opening cost f_j and finite capacity s_j), a set of clients C with variable demands, and a distance metric $d(i, j)$ defined between client i and facility j, the capacitated Facility Location Problem (cFLP) consists in selecting a set S of facilities to open and an assignment of demands to open facilities that minimize the sum of facility opening cost, supplying cost and connection cost (total distance between every client and the facility it is connected to). This problem shares many similarities with the one consisting in minimizing the cost for opening a set of servers or replicas for storing a set of digital objects (data) and connecting clients to facility locations so as to satisfy their demands. The multi-source dimension of the problem translates the situation where the same data object may be available simultaneously at different facility locations. When data objects of different types are available simultaneously at multiple facility locations, the problem shares the characteristics of a multi-product or multi-commodity model. Moreover the dynamics of client demands may lead to consider the replication of data objects at multiple facility locations depending on the capacity allocation and sharing model together with their associated constraints; therefore, client demands can be possibly assigned to locations other than the closest facility. Nevertheless, some specific characteristics of our problem need to be further emphasized as data objects are not physical goods in constrast to classical facility location problems, i.e., servicing one customer doesn't prevent servicing another customer requesting for the same data object.

On the other hand, due to the inherent variability of client demands resulting from many exogeneous dependencies, the input data to the problem is subject to uncertainty. To take this uncertainty into account as part of the problem formulation, we customize robust optimization methods to find solutions which remain valid even if the input data (the spatio-temporal properties of client demands) changes. We further consider the set induced robust optimization method where the uncertain data are assumed to be varying in a given uncertainty set $\mathcal{U}$. The goal is to choose the best solution among those protecting against data uncertainty, i.e., among the candidate solutions that remain feasible for all realizations of the scenarios described by the parameters appearing in the constraints and/or the objective function and belonging to the uncertainty set $\mathcal{U}$. Optimizing the worst-case over all scenarios, referred to as strict robustness contrasts to less conservative methods which overcome usual situations (appearing in practical applications) where the set of strictly robust solutions is actually empty, or all of the strictly robust solutions lead to undesirable solutions (i.e., when yielding inadequate objective function values). Among these less conservative methods, the Γ-robustness method (and its variants) which relies on the working assumption that it is unlikely that all coefficients of one constraint change simultaneously to their worst-case values, limits the uncertainty set $\mathcal{U}$ by fixing a real number for every constraint and protect against the case where that at most Γ coefficients deviate. For this purpose, we further assume that i) the client demands are uncertain for all the commodities, i.e., their value is not known exactly when the optimization problem is solved and ii) data uncertainty can be modeled by means of box+polyhedral uncertainty sets leading to consider client demands within their estimated value and its distributional deviation around that value. On the other hand, dealing with uncertainty in client demands may be considered in the context of both a single-period formulation (where demands may vary within a single period due to the dependence on the requested object) and a multi-period formulation (where demands dynamically vary over time due to their dependence on the client utility, activity, and expectation).

In this paper, we propose and compare different robust formulations (depending on the features of the uncertainty sets) together with resolution methods of the multi-source multi-commodity capacitated facility location problem adapted for replicas placement and data replication. Results of numerical experiments performed on representative settings and running scenarios are reported and analyzed. We further investigate the dependence of the solution properties and the computational time to produce them with respect to the size of the instances under consideration. Finally, we exploit these results to compare different replicas placement and data replication strategies against various capacity allocation and resource sharing models between replicas.