On capacitated clustering problem.
Preprint
DOI: 10.13140/RG.2.2.33224.70401
November 21, 2023

Mark Sh. Levin

The paper addresses capacitated clustering problems: (a) basic capacitated clustering problem, (b) capacitated centered clustering problem, (c) multi-capacity clustering problem, (d) related problems. The paper material is based on combinatorial clustering viewpoint. A survey on the problems, solving approaches, and some applications is presented. The optimization models of basic capacitated clustering problem and multicriteria capacitated clustering problem are considered. Two applications of capacitated clustering in communication networks are briefly described: (a) handover minimization in mobile wireless networks, (b) allocation of end-users to access points in telecommunication networks. Numerical examples illustrate problems and applications.

Keywords: capacitated clustering, combinatorial clustering, combinatorial optimization, communications

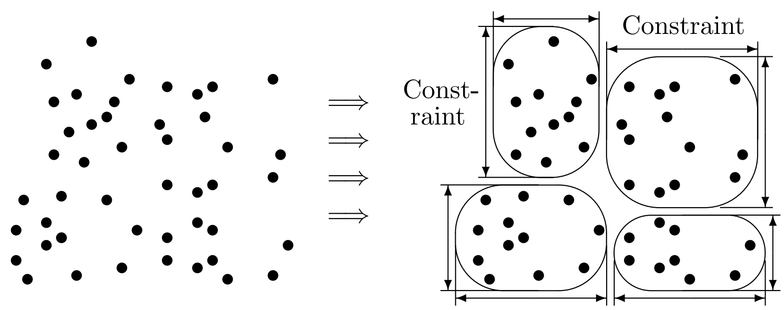
Contents

1. Introduction

In recent years a special attention is targeted to various combinatorial clustering problems and approaches (e.g., [13,27,42,44,45,52,69,88,91]). The paper addresses capacitated clustering problems as a special type of combinatorial clustering (e.g., [4,16,20,34,63,66,67,71]). This problem is also known as the node capacitated graph partitioning problem [24,70]. This material is a part of of combinatorial clustering engineering approach (e.g., [51,52,54–59]).

The capacitated clustering problem (CCP) consists in forming a specified number of clusters (or groups) from a set of elements in such a way that the sum of the weights of the elements in each cluster is within some capacity limits, and the sum of the benefits between the pairs of elements in the same cluster is maximized. CPP is a hard combinatorial optimization problem (e.g., [4,71]) and is widely studied in the last few decades. An illustration of the capacitated clustering problem (4 clusters) is depicted in Fig. 1.

Mainly the following four basic capacitated clustering problems are examined: (a) basic capacitated clustering problem (CCP) (e.g., [20,65–67,71]); (b) capacitated centered clustering problem (CCCP) (e.g., [14,16,34,72,73,92]); (c) multi-capacity clustering problem (MCCP) (e.g., [4,77]); and (d) heterogeneous capacitated clustering problem (HCCP) [74]. Basic capacitated clustering problems and some prospective problems are pointed out in Table 1.



(a) initial set of points (b) capacitated clustering solution

Fig. 1. Illustration for capacitated clustering

Table 1. Capacitated clustering problems

No.	Problem	Source(s)
1.	Capatitated clustering problem formulations:	
1.1.	Survey on the on the problem formulations (and on solving heuristics)	[20]
1.2.	Basic capacitated clustering problem	[63,65–67,71]
1.3.	Capacitated centered clustering	[14–16,34,68,72,73,92]
1.4.	Multi-capacity clustering problem	[4,77]
1.5.	Heterogeneous capacitated clustering problems	[74]
1.6.	Heterogeneous capacitated centered clustering problem	[74]
1.7.	Large-scale capacitated clustering	[34]
1.8.	Capacitated clustering problem on the tree	[33]
1.9.	Fair-capacitated clustering	[79]
2.	Some prospective problem formulations:	
2.1.	Multicriteria capacitated clustering problem	[36], this paper
2.2.	Capacitated clustering problem with ordinal estimates	
2.3.	Capacitated clustering problem under uncertainty (e.g., stochastic problems, with fuzzy set estimates, with multiset estimates)	
2.4.	Multicriteria capacitated centered clustering problem	
2.5.	Capacitated centered clustering problem with ordinal estimates	
2.6.	Capacitated centered clustering problem under uncertainty (e.g., stochastic problems, with fuzzy set estimates, with multiset estimates)	
2.7.	Multicriteria multi-capacitated clustering problem	
2.8.	Multi-capacitated clustering problem with ordinal estimates	
2.9.	Multi-capacitated clustering problem under uncertainty (e.g., stochastic problems, with fuzzy set estimates, with multiset estimates)	

Fig. 2 depicts a framework of main capacity clustering problem types and some related problems/domains.

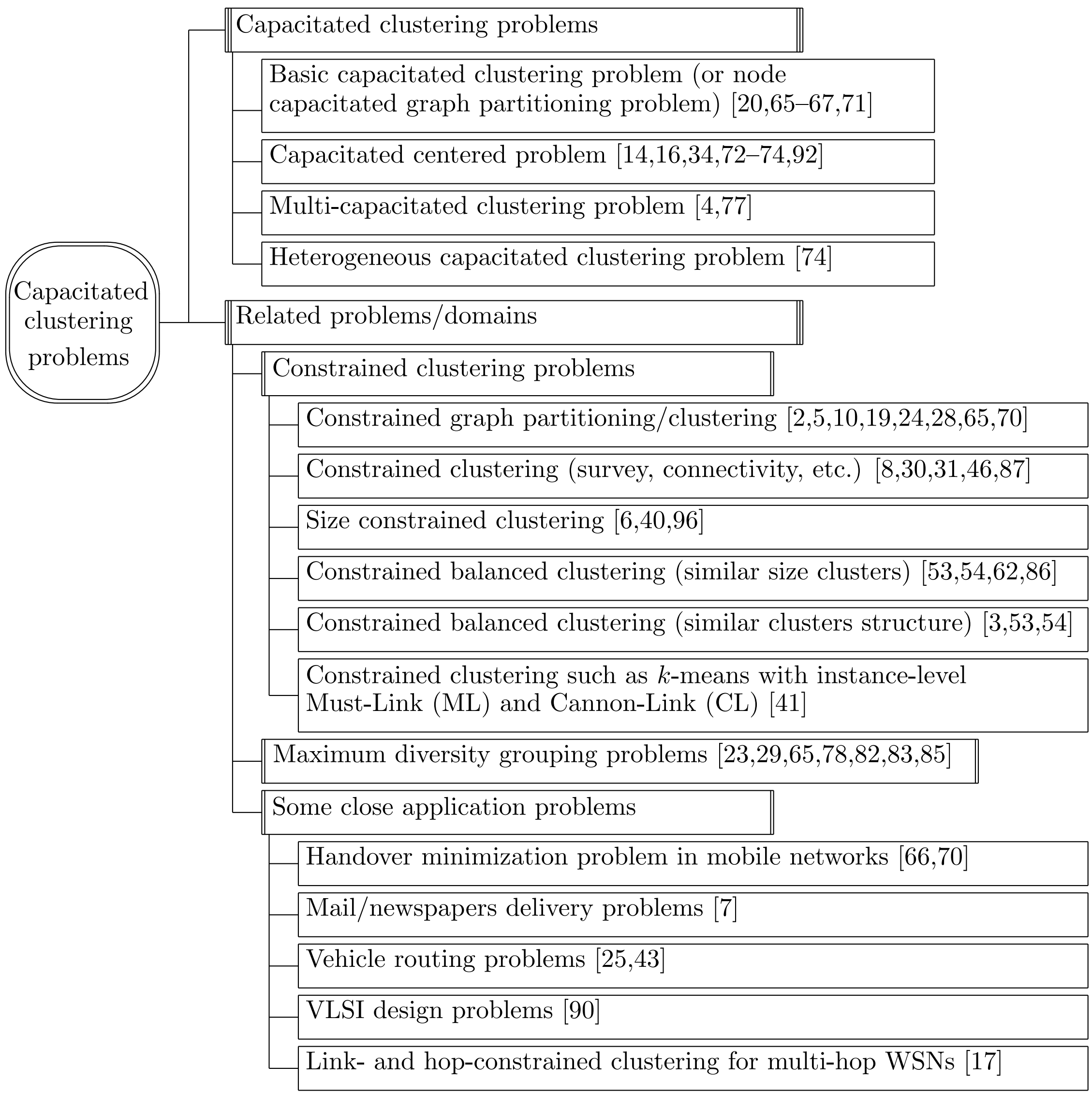


Fig. 2. Capacity clustering problems and related problems/domains

2. Description of basic capacitated clustering problems

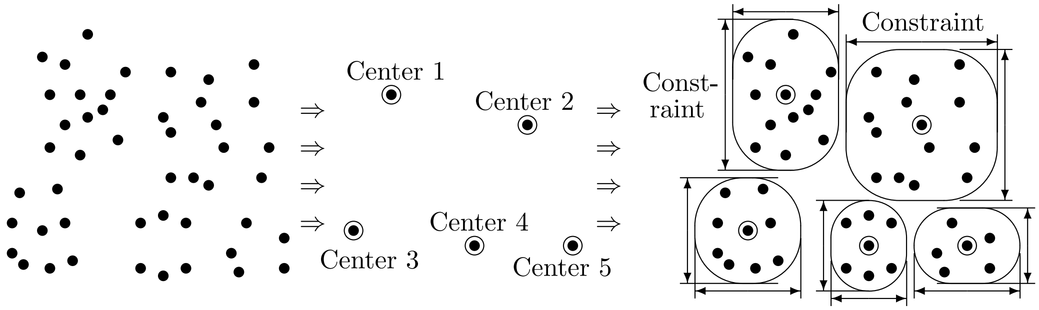
In the basic capacitated clustering problem (CCP) the undirected graph is considered $G = (A, E)$ (A is the vertices/nodes set, E is the edge set), (e.g., [20,63,65,66,71,95]). CCP consists in partitioning the node set A into p disjoint clusters/groups (clustering solution $\widehat{X} = \{X_1, ..., X_k, ..., X_p\}$) in such a way that:

(1) the sum of the weights of the elements (nodes) in each cluster is within some integer capacity limits, Q^-_k and Q^+_k ($Q^-_k < Q^+_k$), and

(2) the sum of the benefits between the pairs of elements (nodes) in the same cluster is maximized.

The basic Capacitated Centered Clustering Problem (CCCP) (e.g., [14,16,34,68,72,73,92]) consists in partitioning a set of n points in a space R^λ with dimension $\lambda \geq 2$ into p disjoint clusters with a known capacity and each cluster is specified by a centroid. The objective is to minimize the total dissimilarity within each cluster (i.e., the sum of the Euclidean distances between the points and their respective cluster centroids), such that a given capacity limit of the cluster is not exceeded. Recently a number of applications of CCCP is described in literature (e.g., in the dry food distribution logistics, in designing zones for urban garbage collection, in territorial design of salesmen regions).

An illustration of the capacitated centered clustering problem (5 centroids corresponding to clusters) is depicted in Fig. 3. Here cluster centers are selected for clusters 1, 2, and 3 and cluster centers are added (defined) for cluster 4 and 5.



(a) initial set of points (b) cluster centers (c) centered capacitated clustering

Fig. 3. Illustration for capacitated centered clustering

The basic CCCP (or p-CCCP) is non-linear. It is NP-Hard, once its unconstrained version is also NP-Hard [39]. A special class of p-center p-supplier problems is under consideration (e.g., [11,38,80]) (Table 2). In fact, the problems are based on capacitated p-center clustering.

Table 2. Constrained p-center p-supplier problems

No.	Problem
1.	p-supplier problem
2.	Contrained p-center and p-supplier problem (tight FPT approximation)
3.	r-gather k-supplier problem (with and without outliers)
4.	r-capacity p-supplier p-center problem (without outliers)
5.	Balanced p-supplier problem (with non-uniform lower and upper bounds)
6.	Chromatic p-supplier problem
7.	Fault-tolerant p-supplier problem
8.	Strongly private p-supplier problem
9.	l-diversity p-supplier problem
10.	Fair p-supplier problem

In the balanced p-center capacitated clustering problem both lower and up constraints on cluster size are examined [21].

The basic multi-capacity clustering problem (MCCP) was proposed in [4,77]. In this problem various types of elements (individuals) are considered. The problem consists in a capacitated clustering problem in which each cluster has a given capacity for each type of elements (individual). If all elements are of the same type, the problem is reduced to the classical capacitated clustering problem. The existence of multiple types of individuals, as well as multiple capacities in the facilities, increases the complexity of this new variant of the clustering problem. MCCP is different from capacitated p-median and capacitated cluster problems since the capacities may be different between types and medians.

A simplified illustration for multi-capacitated clustering (here: clustering with multi-type elements) is depicted in Fig. 4 (four clusters are designed on the basis of four elements types):

1. cluster 1: 6 elements (4 of type 1, 1 element of type 3, 1 element of type 4);
2. cluster 2: 7 elements (5 of type 2, 1 of type 1, 1 of type 4);
3. cluster 3: 8 elements (1 element of type 1, 6 of type 3, 1 of type 4); and
4. cluster 4: 8 elements (1 element of type 1, 1 element of type 2, 6 elements of type 4).

Evidently, two constraints for each element type have to be considered as well.

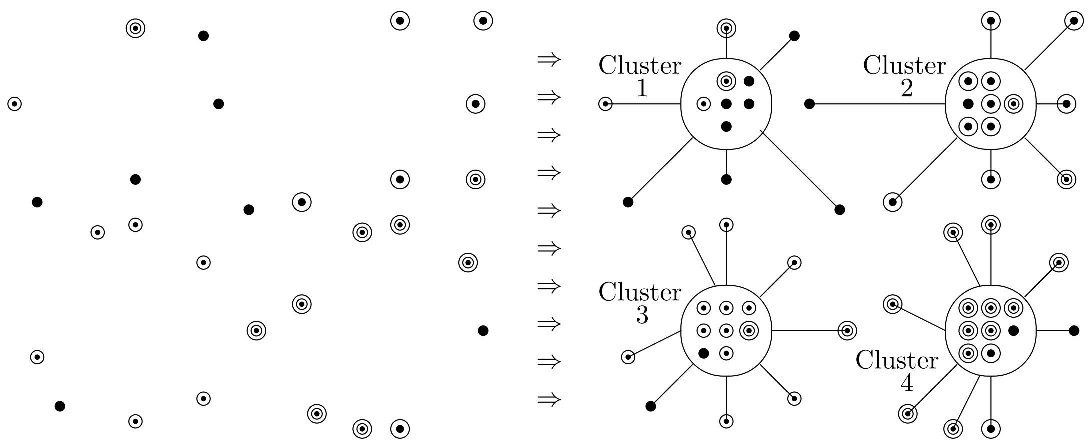


Fig. 4. Illustrative example of multi-capacitated clustering

3. Related problems

3.1. Some basic related problems

The following basic related problems are pointed out in the literature (Fig. 2):

1. Capacitated clustering problem (CCP) is closely related to the graph partition problem (GPP) (e.g., [4,95]), where the goal is to find a partition of the vertex set in p classes while minimizing the number of cut edges and respecting a balance constraint between the classes.

2. Various constrained clustering problems, for example: (i) constrained graph partitioning/clustering (e.g., [2,5,10,19,24,28,65,70]); (ii) size constrained clustering (e.g., [6,40,96]); (iii) constrained balanced clustering (e.g., [3,53,54,62,86]).

3. The capacitated clustering problem is close to facility constrained location problem (e.g., [50,60,61]).

4. Capacitated clustering problem (CCP) is equivalent to the Handover Minimization Problem (HMP) in mobility networks (e.g., [66,70]). Here the objective is to minimize the sum of edge weights between elements of different clusters.

5. Another important related problem is the maximum diversity (or maximally diverse) grouping problem (MDGP) (e.g., [49,65,82,83,85,94,95]):

Some real-world applications of capacitated clustering problems are pointed out later.

3.2. Maximum diversity grouping problem

An interesting close problem is considered as maximum diversity grouping problem (MDGP) (e.g., [23,29,49,65,78,82,83,85,94]):

Find the assignment of a set of items ($i \in A$) to disjoint groups in such a way that the diversity among the elements in each group (the sum of pairwise distances/benefits c_{ij} between all items assigned to the same group) is maximized.

Here the heterogeneity inside groups is maximized. The objective of the problem is to maximize the overall diversity, i.e., the sum of the diversity of all groups, when the size of each group is within a specified range. Clearly, the MDGP is a special case of the CCP.

The MDGP is called the p-partition problem in [23] and the equitable partition problem in [75]. The problem belongs to the family of diversity problems (e.g., [22,29,78]).

3.3. Some real-world applications

Some examples of real-world applications of capacitated clustering in industry and services which are described in literature are listed in Table 3 (e.g., [26,32,64,71,73]).

Table 3. Real-world applications of capacitated clustering in industry and services

No.	Application	Source(s)
1.	Hop-constrained clustering for multi-hop WSNs	[17]
2.	Partitioning of nodes in distributed communication networks	[4]
3.	Healthcare application (e.g., assignment of people to a certain hospital(s))	[4]
4.	Vehicle routing, consolidation of customer orders into vehicle shipments	[25,43,95]
5.	Designing zones for urban garbage collection	[73]
6.	Mail delivery, newspaper delivery	[7,32,71,73]
7.	Layout of IT-teams in software factories	[74]
8.	VLSI design	[64,95]
9.	Educational applications (forming students groups/teams)	[4,51,53,54,89]
10.	Products manufacturing management (e.g., clustering of products in facilities like refinery)	[4]
11.	Oil industry (assignment of the oil by-products demands among the existing refineries facilities)	[4]

4. Basic solving approaches

Table 4 contains a list of basic solving approaches.

5. Some formulations of capacitated clustering problems

5.1. Basic capacitated clustering problem

The following components are examined in the basic capacitated clustering problem (e.g., [20,63,65,66, 71,95]):

(a) graph $G = (A, E)$ where A is a set of n nodes (vertices, elements, items) (i.e., $A = \{1, ..., i, ..., n\}$ and E is a set of edges $E = \{(i_1, i_2)\}, \ \forall i_1, i_2 \in A$;

(b) node weight (e.g., demand) $w_i \geq 0$ ($\forall i \in A$);

(c) benefit (or weight/profit/utility) of edge c_{ij} ($\forall (i,j) \in E$), note $c_{ij} = 0$ if edge $(i,j) \overline{\in} E$;

(d) given number p of disjoint clusters $\{X_1 = (A_1, E_1), ..., X_k = (A_k, E_k), ..., X_p = (A_p, E_p)\}$ (i.e., $k = \overline{1,p}$, here $|A_{k_1} \& A_{k_2}| = 0 \quad \forall k_1, k_2 = \overline{1,k}$ and $k_1 \neq k_2$);

(e) two constraints are considered for each cluster $k = \overline{1,p}$: (i) minimum capacity (i.e., constraint) Q^-_k and (ii) maximum capacity (i.e., constraint) Q^+_k.

The capacitated clustering problem (CCP) is usually formulated as the following quadratic integer program (e.g., [12,20,95]) with binary variables: $x_{ik} = 1$ if element (node) i is in cluster k and 0 otherwise. Thus the basic optimization model is:

$$\max \ \sum_{k=1}^{p} \sum_{i=1}^{n-1} \sum_{j>i}^{n} c_{ij} \ x_{ik} \ x_{jk} \qquad (1.1)$$

$$s.t. \quad \sum_{k=1}^{p} x_{ik} = 1, \quad i = \overline{1,n}, \qquad (1.2)$$

$$Q^-_k \leq \sum_{i=1}^{n} w_i x_{ik} \leq Q^+_k, \quad k = \overline{1,p}, \qquad (1.3)$$

$$x_{ik} \in \{0,1\}, \quad i = \overline{1,n}, \quad k = \overline{1,p} \qquad (1.4)$$

Here the objective function (1.1) corresponds to the total benefit of all pairs of elements that belong to the same cluster.

Table 4. Basic solving approaches

No.	Approach	Source(s)
1.	Surveys, generalized algorithm descriptions:	
1.1.	Comprehensive review on the most representative approaches for CCP	[20]
1.2.	Bibliographic view (survey)	[46]
1.3.	Survey on state-of-the-art approaches (heuristic search, greedy randomize adaptive search procedures (GRASP), etc.)	[95]
1.4.	State-of-the-art heuristics for capacitated clustering problem (Tabu search, randomized algorithms, hybrid methods)	[65,66]
1.5.	Pairwise-confidence-constraints-clustering algorithm (center-based heuristic)	[9]
2.	Heuristics:	
2.1.	Hybrid simulated annealing and tabu search (for capacitated clustering problems)	[76]
2.2.	Scatter search heuristic for capacitated clustering problem	[81]
2.3.	Tabu search for the capacitated clustering problem	[95]
2.4.	Path-relinking with tabu search for the capacitated centered clustering problem	[72]
2.5.	Tabu search and GRASP for the capacitated clustering problem	[65]
2.6.	Randomized heuristics for the capacitated clustering problem	[66]
2.7.	Greedy random adaptive memory programming search for capacitated clustering	[1]
3.	Some special search based methods (e.g., local search, GRASP, etc.):	
3.1.	Adaptive biased random-key GA with local search for capacitated centered clustering	[16]
3.2.	Parallel clustering search applied to capacitated centered clustering problem	[68]
3.3.	Constrained clustering through dual iterative local search (new metaheuristic)	[35]
3.4.	Reactive GRASP (Greedy Randomized Adaptive Search Procedure) with path relinking for capacitated clustering	[20]
4.	Variable neighborhood search (VNS) approaches:	
4.1.	Variable neighborhood search (VNS) for capacitated clustering problem	[12,48]
4.2.	Iterated VNS for the capacitated clustering problem	[47]
4.3.	Neighborhood decomposition-driven VNS for capacitated clustering	[48]
4.4.	Iterative neighborhood local search algorithm for capacitated centered clustering	[92]
5.	Evolutionary methods:	
5.1.	Genetic algorithms for capacitated clustering problem	[84]
5.2.	Hybrid evolutionary algorithm for the capacitated centered clustering problem	[15]
5.3.	Differential evolution approach for instance-level constrained clustering	[37]
5.4.	Tabu search memetic algorithm for the capacitated clustering problem	[95]
5.5.	Membrane evolutionary algorithm for capacitated clustering problem (MEACCP)	[64]
5.6.	Decomposition-based memetic elitism (for multiobjective constrained clustering)	[36]
6.	Metaheurisrics:	
6.1.	Metaheuristic framework for heterogeneous capacitated centered clustering (HCCCP)	[74]
6.2.	Clustering search algorithm as hybrid metaheuristic for capacitated centered clustering	[14]
6.3.	Hybrid metaheuristics for multi-capacity clustering problem	[4]
6.4.	Matheuristic for large-scale capacitated clustering	[34]
7.	Hybrid methods:	
7.1.	HA-CCP: a hybrid algorithm for solving capacitated clustering problem	[63]
7.2.	Hybrid metaheuristics for multi-capacity clustering problem	[4]
7.3.	Hybrid evolutionary algorithm for the capacitated centered clustering problem	[15]
7.4.	Clustering search algorithm as hybrid metaheuristic for capacitated centered clustering problem	[14]
7.5.	Three-phase search approach with dynamic population size (hybrid algorithm) (for solving the maximally diverse grouping problem)	[94]
8.	Some special approaches:	
8.1.	Lagrangian relaxation approach for a large scale new variant of capacitated clustering problem	[93]
8.2.	Fixed-parameter tractability of capacitated clustering	[18]

In the model above the constraints are as follows:

(i) the constraint (1.2) corresponds the assignment of each element to a cluster (i.e., the constraint guarantees that every node is assigned to exactly one cluster), and

(ii) the constraint (1.3) corresponds to the following: the sum of the weights of the elements in the same cluster k ($k = \overline{1,p}$) to be between the specified constraints: Q^-_k and Q^+_k (i.e, the constraint ensures that the minimum capacity Q^-_k and the maximum capacity Q^+_k requirements of each cluster k are satisfied). Evidently in a simplified problem the same capacity constraints can be used for each cluster k: $Q^-_k = Q^-$ ($k = \overline{1,p}$) and $Q^+_k = Q^+$ ($k = \overline{1,p}$).

5.2. Multicriteria capacitated clustering problem

First, in multicriteria capacitated clustering problem the benefit (profit/utility) of edge $(i,j) \in E$ $c_{i,j}$ can be transformed into the vector $\overline{c}_{i,j} = (c^1_{i,j}, ..., c^\gamma_{i,j}, ..., c^\mu_{i,j})$. The corresponding problem will be as follows:

$$\max \sum_{k=1}^{p} \sum_{i=1}^{n-1} \sum_{j>i}^{n} c^1_{ij} x_{ik}\ x_{jk}, ..., \max \sum_{k=1}^{p} \sum_{i=1}^{n-1} \sum_{j>i}^{n} c^\gamma_{ij} x_{ik}\ x_{jk}, ..., \max \sum_{k=1}^{p} \sum_{i=1}^{n-1} \sum_{j>i}^{n} c^\mu_{ij} x_{ik}\ x_{jk}; \quad (2.1)$$

$$s.t. \quad \sum_{k=1}^{p} x_{ik} = 1, \quad i = \overline{1,n} \quad (2.2)$$

$$Q^-_k \le \sum_{i=1}^{n} w_i x_{ik} \le Q^+_k, \quad k = \overline{1,p}, \quad (2.3)$$

$$x_{ik} \in \{0,1\}, \quad i = \overline{1,n}, \quad k = \overline{1,p} \quad (2.4)$$

Here the Pareto efficient solution(s) have to be search for.

Second, the node weight w_i ($i \in A$) can be transformed into vector weight $\overline{w}_i = (w^1_i, ...w^\xi_i, ...w^\eta_i)$ and corresponding constraint for each cluster $k \in \{1, ..., p\}$ (i.e., constraint (2.3) will be changed by a set of η constraints. Thus the model is:

$$\max \sum_{k=1}^{p} \sum_{i=1}^{n-1} \sum_{j>i}^{n} c^1_{ij} x_{ik}\ x_{jk}, ..., \max \sum_{k=1}^{p} \sum_{i=1}^{n-1} \sum_{j>i}^{n} c^\gamma_{ij} x_{ik}\ x_{jk}, ..., \max \sum_{k=1}^{p} \sum_{i=1}^{n-1} \sum_{j>i}^{n} c^\mu_{ij} x_{ik}\ x_{jk}; \quad (3.1)$$

$$s.t. \quad \sum_{k=1}^{p} x_{ik} = 1, \quad i = \overline{1,n} \quad (3.2)$$

$$Q^{\xi -}_k \le \sum_{i=1}^{n} w^\xi_i x_{ik} \le Q^{\xi +}_k, \ k = \overline{1,p}, \ \xi = \overline{1,\eta}, \quad (3.3)$$

$$x_{ik} \in \{0,1\}, \quad i = \overline{1,n}, \quad k = \overline{1,p} \quad (3.4)$$

Evidently various modifications of multicriteria capacitated clustering problem can be examined as well.

6. Applied examples of capacitated clustering in communications

In general, main clustering problems in networking are targeted to design some clusters under special constraints (e.g., resource constraints). In this section two special applied clustering problems on communication networks are briefly described (at an illustrative level).

6.1. Handover minimization in mobile wireless networks

In recent years, the application of capacitated clustering problem (or node capacitated graph partitioning problem) for handover minimization in mobile wireless networks is examined (e.g., [66,70]). An illustrative example of mobility network is depicted in Fig. 5: mobile users (end users), base stations, and radio network controllers (RNC). Radio network controllers are controlling the base station operations, including traffic and handover. Handovers between base stations connected to different RNCs tend to fail more often than handovers between base stations connected to the same RNC. The Handover Minimization Problem consists in assignment of the base stations to RNCs. A subset of base stations which are assigned to the same RNS can be considered as a cluster. The minimization of handovers between different clusters is equivalent to the maximization of handovers within the same cluster.

Here the following notations are used: (i) set of base stations A = {1,...,i,...,n}, (ii) graph $G = (A, E)$, E is a set of edges (i.e., pair connections between base stations) in G, (iii) edge benefit c_u ($\forall u \in E$), (iv) node weight $w_i \geq 0$ ($\forall i \in A$), (v) integer capacity limits for each cluster Q^+ and Q^+. The problem is equivalent to the CCP:

Find a partition of set A (base stations) into p clusters (groups of base stations – each group/cluster is assigned to the same controller) such that:

(a) the sum of edge benefits (benefit c_u of edge $u \in E$) in clusters is maximized, and

(b) for each cluster the sum of the node weights (w_i, for nodes $i \in A$) satisfied to capacity limits: Q^- and Q^+ ($Q^- < Q^+$).

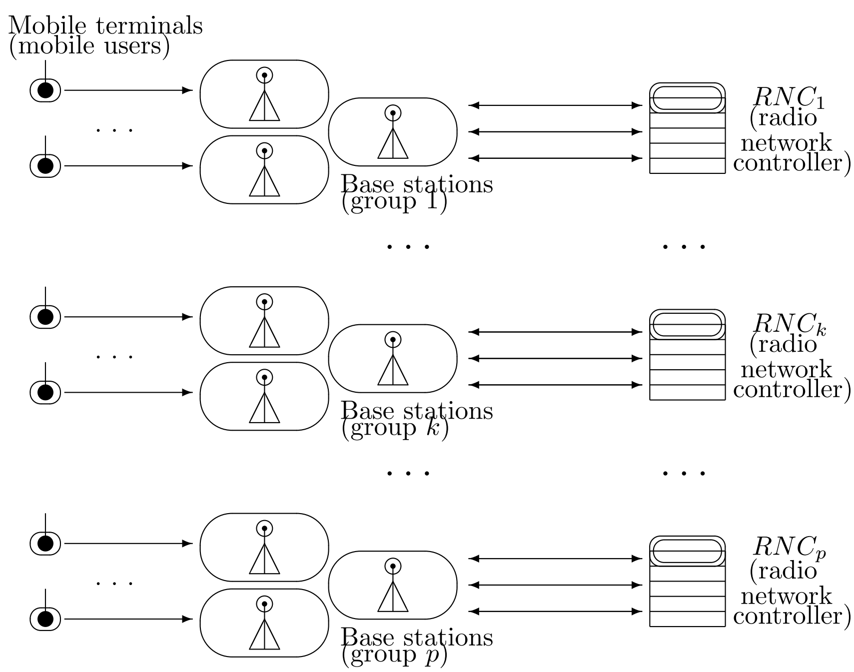


Fig. 5. Illustration of handover minimization in mobile wireless networks

6.2. Connection of users with telecommunication networks

A special multicriteria assignment applied problem is examined in [60,61]. Here a communication network involves a set of end-users and a set of access points. A set of end-users are divided into groups (clusters) and for each cluster the corresponding end-users are connected with the same access point of the communication network while taking into account the following: (a) a resource constraint of each access point, (b) proximities (e.g., distance) between end-points and access points, (c) additional parameters of the connection between the end-user and the access point (e.g., reliability of the connection) and (d) end-user requirement(s) for information transmission (e.g., frequency bandwidth, required level of connection reliability).

Thus the problem consists in allocation of end-users to access points (i.e., grouping/clustering of end-users via multicriteria assignment and resource constraints). Clearly, this problem is very close to capacitated clustering problem and can be considered as a special modification of the capacitated clustering. A simplified illustrative numerical example of the problem (i.e., clustering solution) is depicted in Fig. 6 (set of end-users $A = \{1, 2, ..., 24\}$, (5 access points $B = \{1, 2, 3, 4, 5\}$):

(1) cluster 1 (access point 1): $X_1 = \{1, 2, 3, 4, 5, 6\}$,
(2) cluster 2 (access point 2): $X_2 = \{7, 8, 9, 10, 11\}$,
(3) cluster 3 (access point 3): $X_3 = \{12, 13, 14\}$,
(4) cluster 4 (access point 4): $X_4 = \{15, 16, 17, 18, 19\}$,
(5) cluster 5 (access point 5): $X_5 = \{20, 21, 22, 23, 24\}$.

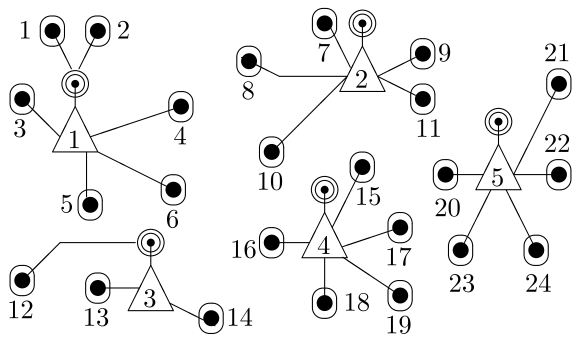


Fig. 6. Assignment of users to access points

7. Conclusion

In the paper versions of capacitated clustering problems are examined from the viewpoint of combinatorial clustering: (a) basic capacitated clustering problem, (b) capacitated centered clustering problem, (b) multi-capacity clustering problem, (c) some real-world related problems (including maximum diversity grouping problem). A survey on the problems, solving approaches, and some applications is presented. The optimization formulations of basic capacitated clustering problem and two versions of multicriteria capacitated clustering problem are considered. Two network applications of capacitated clustering are briefly described (at an illustration level): (a) handover minimization in mobile wireless networks, (b) allocation of end-users to access points in telecommunication networks.

In the future, it may be reasonable to consider the following: (1) examination of new multicriteria models for capacitated centered clustering problem and multi-capacity clustering problem; (2) special study of various versions of the capacitated clustering problem under uncertainty (stochastic statement, models with fuzzy estimates, models with multiset estimates, etc.); (3) design and analysis of special new solving strategies (e.g., special metaheuristics and hybrid solving schemes); (4) special examination of maximum diversity grouping problem and its applications; (5) additional description and analysis of capacitated clustering problems in various application domains; and (6) using the suggested problems and solving frameworks in education.

The author states that there is no conflict of interest.

REFERENCES

1. S. Ahmadi, I.H. Osman, Greedy random adaptive memory programming search for the capacitated clustering problems. EJOR, 162, 30-44, 2005.
2. R. Anand, C.K. Reddy, Graph-based clustering with constraints. In: J.Z. Huang, L. Cao, J. Srivastava (eds), Proc. of 15th Pacific-Asia Conf. on Advances in Knowledge Discovery and Data Mining (PAKDD 2011), Part II, LNCS 6635, Springer, pp. 51–62, 2011.
3. S.A.M. Anaraki, A. Haeri, Soft and hard hybrid balanced clustering with innovative qualitative balancing approach. Inf. Sci., 613, 786–805, 2022.
4. K.A.G. de Araujo, J.B. Guedes, B. de Athayde Prata, Hybrid metaheuristics for the multi-capacitated clustering problem. RAIRO-Oper. Res., 56, 1167-1185, 2022.
5. A. Banerjee, J. Ghosh, Scalable clustering algorithms with balancing constraints. Data Mining and Knowledge Discovery, 13, 365-395, 2006.
6. M. Baranwal, S.M. Salapaka, Clustering with capacity and size constraints: a deterministic approach. In: 2017 Indian Control Conf. (ICC), pp. 251-256, 2017.

7. J.F. Bard, A.I. Jarrah, Large-scale constrained clustering for rationalizing pickup and delivery operations. Transportation Research Part B: Methodological, 43(5), 542–561, 2009.
8. S. Basu, I. Davidson, K. Wagstaff, Constrained Clustering: Advances in Algorithms, Theory and Applications. CRC Press, 2008.
9. P. Baumann, D.S. Hochbaum, PCCC: the pairwise-confidence-constraints-clustering algorithm. Electr. prepr., 79 p., Jan. 12, 2023. http://arxiv.org/abs/2212.14437 [cs.LG]
10. U. Benlic, J.K. Hao, A multilevel memetic approach for improving graph k-partitions. IEEE Trans. on Evolutionary Computation, 15(5), 624–642, 2011.
11. I.O. Bercea, M. Gross, S. Khuller, A. Kumar, C. Rosner, D.R. Schmidt, M. Schmidt, On the cost of essentially fair clustering. In: D. Achlioptas, L.A. Vegh (eds), Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques APPROX/Random 2019, pp. 18:1–18:22, 2019.
12. J. Brimberg, N. Mladenovic, R. Todosijevic, D. Urosevic, Solving the capacitated clustering problem with variable neighborhood search. Annals of Oper. Res., 272(1), 289-321, 2018.
13. Y. Cao, Z. Wang,Combinatorial optimization-based clustering algorithm for wireless sensor networks. Mathematical Problems in Engineering, art. ID 613704, 1–13, 2020.
14. A.A. Chaves, L. Lorena, Clustering search algorithm for the capacitated centered clustering problem. Comp. & Oper. Res., 37(3), 552–558, 2010.
15. A.A. Chaves, L. Lorena, Hybrid evolutionary algorithm for the capacitated centered clustering problem. ESwA, 38, 5013–5018, 2011.
16. A.A. Chaves, J.F. Goncalves, L.A.N. Lorena, Adaptive biased random-key genetic algorithm with local search for the capacitated centered clustering problem. Comp. & Ind. Eng., 124, 331–346, 2018.
17. D.-R. Chen, A link- and hop-constrained clustering for multi-hop wireless sensor networks. Comp. Commun., 72, 78–82, 2015.
18. V. Cohen-Addad, J. Li, On the fixed-parameter tractability of capacitated clustering. In: 42nd Conf. on Very Important Topics (CVIT 2016), art. no. 23, pp. 23:1–23:15, 2016.
19. V.-T. Dang, V.-V. Vu, H.-Q. Do, T.K.O. Le, Graph-based clustering with constraints and active learning. J. of Computer Science and Cybernetics, 37(1), 71–89, 2021.
20. Y. Deng, J.F. Bard, A reactive GRASP with path relinking for capacitated clustering. J. of Heuristics, 17(2), 119–152, 2011.
21. H. Ding, Balanced k-center clustering when k is A constant. Electr. prepr., 6 p., Apr. 8, 2017. http://arxiv.org/abs/1704.02515 [cs.CG]
22. A. Duarte, J. Snchez-Oro, M. Resende, F. Glover, R. Mart, Greedy randomized adaptive search procedure with exterior path relinking for differential dispersion minimization. Inform. Sci., 296, 46-60, 2015.
23. T. Feo, O. Goldschmidt, M. Khellaf, One-half approximation algorithms for the k-partition problem. Oper. Res. 40, S170-S173, 1992.
24. C.E. Ferreira, A. Martin, C.C. de Souza, R. Weismantel, L.A. Wolsey, The node capacitated graph partitioning problem: a computational study. Math. Program., 81, 229-256, 1998.
25. M.L. Fisher, R. Jaikumar, A generalized assignment heuristic for vehicle routing. Networks, 11(2), 109–124, 1981.
26. P.M. Franca, N.M. Sosa, V. Pureza, An adaptive tabu search algorithm for the capacitated clustering problem. Int. Trans. Oper. Res., 6, 665-678, 1999.
27. N.J. Fuda, K. Brejc, W.S. Kruesi, E.J. Ralston, R. Bigley, A. Shin, M. Okada, B.J. Meyer, Combinatorial clustering of distinct DNA motifs directs synergistic binding of *Caenorhabditis elegans* dosage compensation complex to X chromosomes. PNAS, 119(37), e2211642119, 1–9, 2022.
28. P. Galinier, Z. Boujbel, M.C. Fernandes, An efficient memetic algorithm for the graph partitioning problem. Ann. of Oper. Res., 191(1), 1–22, 2011.
29. M. Gallego, A. Duarte, M. Laguna, R. Mart, Hybrid heuristics for the maximum diversity problem. Comput. Optim. Appl., 44(3), 411–426, 2009.
30. P. Gancarski, T.-B.-H. Dao, B. Cremilleux, G. Forestier, T. Lampert, Constrained clustering: current and new trends. In: P. Marquis, P. Papini, H. Prade (eds), A Guided Tour of Artificial Intelligence Research, Springer, pp. 447–484, 2020.
31. J. Gao, X. Tao, S. Cai, Towards more efficient local search algorithms for constrained clustering. Inf.

Sci., 621, 287–307, 2023.
32. S. Geetha, G. Poonthalir, P.T. Vanathi, Improved k-means algorithm for capacitated clustering problem. INFOCOMP J. Comput. Sci., 8, 52-59, 2009.
33. E.Kh. Gimadi, A.A. Kurochkina, E. Nagornaya, On some effective algorithms for solving capacitated clustering and location problems on the tree and the real line. Applications. Irkutsk: ESI SB RAS, p. 95, 2017,
34. M. Gnagi, P. Baumann, A matheuristic for large-scale capacitated clustering. Comp. and Oper. Res., 132, art. 105304, 1–15, 2021.
35. G. Gonzalez-Almagro, J. Luengo, J.-R. Cano, S. Garcia, DILS: constrained clustering through dual iterative local search. Comput. and Oper. Res., 121, art. 104979, 2020.
36. G. Gonzalez-Almagro, A. Rosales-Perez, J. Luengo, J.-R. Cano, S. Garcia, ME-MEOA/D$_{CC}$: multiobjective constrained clustering through decomposition-based memetic elitism. Swarm and Evolutionary Computation, 66, art. 100939, 2020.
37. G. Gonzalez-Almagro, J. Luengo, J.-R. Cano, S. Garcia, Enhancing instance-level constrained clustering through differential evolution. Appl. Soft Comput., 108, art. 107435, 2021.
38. D. Goyal, R. Jaiswal, Tight FPT approximation for constrained k-center and k-supplier. Electr. prepr., 30 p., Oct. 27, 2021. http://arxiv.org/abs/2110.14242 [cs.DS]
39. P. Hansen, B. Jaumard, Cluster analysis and mathematical programming. Mathematical Programming, 79(1), 191-221, 1997.
40. C.W. Hu, H. Li, A.A. Qutub, Shrinkage clustering: a fast and size-constrained clustering algorithm for biomedical applications. BMC Bioinformatics, 19(1), art. 19, 1–11, 2018.
41. C. Jia, L. Guo, K. Liao, Z. Lu, Efficient algorithm for the k-means problem with must-link and cannot-link constraints. Tsinghua Science & Technology, 28(6), 1050-1062, 2023.
42. J. Kim, W. Lee, J.J. Song, S.B. Lee, Optimized combinatorial clustering for stochastic processes. Cluster Computing, 20(2), 1135–1148, 2017.
43. Y.A. Koskosidis, W.B. Powell, Clustering algorithms for consolidation of customer orders into vehicle shipments. Transportation Research, Part B: Methodological, 26(5), 365–379, 1992.
44. M. Kumagai, K. Komatsu, F. Takano, T. Araki, M. Sato, H. Kobayashi, An external definition of the one-hot constraint and fast QUBO generation for high-performance combinatorial clustering. Int. J. of Networking and Computing, 11(2), 463–491, 2021.
45. V. Kumar, G. Bass, C. Tomlin, J. Dulny, Quantum annealing for combinatorial clustering. Quantum Information Processing, 17, 1–14, 2018.
46. L.I. Kuncheva, F.J. Williams, S.L. Hennessey, A bibliographic view on constrained clustering. Electr. prepr., 18 p., Sep. 22, 2022. http://arxiv.org/abs/2209.11125 [cs.LG]
47. X. Lai, J.-K. Hao, Iterated variable neighborhood search for the capacitated clustering problem. Eng. Appl. of Artif. Intell., 56, 102–120, 2016.
48. X. Lai, J.-K. Hao, Z.-H. Fu, D. Yue, Neighborhood decomposition-driven variable neighborhood search for capacitated clustering. Comp. and Oper. Res., 134, art. 105362, 2021.
49. X. Lai, J.-K. Hao, Z.-H. Fu, D. Yue, Neighborhood decomposition based variable neighborhood search and tabu search for maximally diverse grouping. EJOR, 289(3), 1067–1086, 2021.
50. M.Sh. Levin, Modular System Design and Evaluation. Springer, 2015.
51. M.Sh. Levin, Towards combinatorial clustering: preliminary research survey. Electr. prepr., 102 p., May 28, 2015. http://arxiv.org/abs/1505.07872 [cs.AI]
52. M.Sh. Levin, On combinatorial clustering: literature review, methods, examples. J. of Commun. Technol. and Electronics, 60(12), 1403–1428, 2015.
53. M.Sh. Levin, Towards balanced clustering - part 1 (preliminaries). Elec. prepr., 21 p., Jun. 9, 2017. http://arxiv.org/abs/1706.03065 [cs.DS]
54. M.Sh. Levin, On balanced clustering (indices, models, examples). J. of Communications Technology and Electronics, 62(12), 1506–1515, 2017.
55. M.Sh. Levin, Note on dominating sets problems. J. of Communications Technology and Electronics, 66(Suppl. 1), S8–S22, 2021.
56. M.Sh. Levin, Balanced clustering with a tree over clusters. J. of Communications Technology and Electronics, 66(Suppl. 1), S23–S34, 2021.
57. M.Sh. Levin, Clustering models based on graph edge coloring. J. of Communications Technology and

Electronics, 67(12), 1570–1577, 2022.
58. M.Sh. Levin, Towards clique partitioning problem. Preprint, 10 p., Nov. 29, 2022. DOI: 10.13140/RG.2.2.35252.12169 (ResearchGate)
59. M.Sh. Levin, On the clique partitioning of a graph. J. of Communications Technology and Electronics, 67(S2), S267–S274, 2022.
60. M.Sh. Levin, M.V. Petukhov, Multicriteria assignment problem (selection of access points). In: N. Garcia-Pedrajas et al. (eds.), Proc. of 23rd Int. Conf. IEA/AIE 2010, LNCS 6097, part II, Springer, pp. 277–287, 2010.
61. M.Sh. Levin, M.V. Petukhov, Connection of users with telecommunications network: Multicriteria assignment problem. J. of Telecommunications Technology and Electronics 55(12), 1532–1541, 2010.
62. Y. Lin, H. Tang, Y. Li, C. Fang, Z. Xu, Y. Zhou, Z. Zhou, Generating clusters of similar sizes by constrained balanced clustering. Appl. Intell., 52, 5273–5289, 2022.
63. Y. Liu, P. Guo, Y. Zeng, HA-CCP: a hybrid algorithm for solving capacitated clustering problem. Comput. Intell. Neurosci., Volume 2022, art. 6400318, 1–24, 2022
64. Y. Liu, P. Guo, Y. Zeng, MEACCP: A membrane evolutionary algorithm for capacitated clustering problem. Inform. Sci., 591, 319–343, 2022.
65. A. Martinez-Gavara, V. Campos, M. Gallero, M. Laguna, R. Marti, Tabu search and GRASP for the capacitated clustering problem. Computat. Optim. and Appl., 62(2), 589–607, 2015.
66. A. Martinez-Gavara, D. Landa-Silva, V. Campos, R. Marti, Randomized heuristics for the capacitated clustering problem. Information Sciences, 417, 154–168, 2017.
67. A. Mehrotra, M.A. Trick, Cliques and clustering: a combinatorial approach. Oper. Res. Lett., 22, 1–12, 1998.
68. D. Melo Morales, A.A. Chaves, A.L. Fazenda, Parallel clustering search applied to capacitated centered clustering problem. In: 2019 IEEE Int. Parallel and Distributed Processing Symposium Workshops (IPDPSW), pp. 542–548, 2019.
69. B. Mirkin, I. Muchnik, Combinatorial optimization in clustering. In: D.-Z. Du, P.M. Pardalos (eds), Handbook of Combinatorial Optimization. vol. 2, Springer, pp. 261–329, 1999.
70. L.F. Moran-Mirabal, J.L. Gonzalez-Velarde, M.G. Resende, R.M. Silva, Randomized heuristics for handover minimization in mobility networks. J. of Heuristics, 19(6), 845–880, 2013.
71. J.M. Mulvey, M.P. Beck, Solving capacitated clustering problems. EJOR, 18, 339-348, 1984.
72. A.E.F. Muritiba, M.J.N. Gomes, M.F. de Souza, H.L.G. Oria, Path-relinking with tabu search for the capacitated centered clustering problem. ESwA, 198, art. 116766, 2022.
73. M. Negreiros, A. Palhano, The capacitated centred clustering problem. Comput. and Oper. Res., 33, 1639-1663, 2006.
74. M.J. Negreiros, N. Maculan, P.L. Batista, J.A. Rodrigues, A.W.C. Palhano, Capacitated clustering problems applied to the layut of the IT-teams in software factories. Ann. of OR, 316(2), 1157–1185, 2022.
75. F.A. O'Brien, J. Mingers, The Equitable Partitioning Problem: A Heuristic Algorithm Applied to the Allocation of University Student Accommodation. Research Paper No. 187. Warwick Business School, Coventry, 1995.
76. I.H. Osman, N. Christofides, Capacitated clustering problems by hybrid simulated annealing and tabu search. Int. Trans. Oper. Res., 1, 317-336, 1994.
77. B.A. Prata, The multi capacitated clustering problem. Technical report, Federal University of Ceara, Brazil, 2015.
78. O.A. Prokopyev, N. Kong, D.L. Martinez-Torres, The equitable dispersion problem. EJOR, 197, 59-67, 2009.
79. T.L. Quy, A. Roy, G. Friege, E. Ntoutsi, Fair-capacitated clustering. Electr. prepr., 10 p., Apr. 28, 2021. http://arxiv.org/abs/2104.012116 [cs.LG]
80. C. Rosner, M. Schmidt, Privacy preserving clustering with constraints. In: I. Chatzigiannakis, C. Kaklamanis, D. Marx, D. Sannella (eds), Int. Colloquium on Automata, Languages, and Programming ICALP 2018, pp. 96:1–96:14, 2018.
81. S. Scheuerer, R. Wendolsky, A scatter search heuristic for the capacitated clustering problem. EJOR, 169, 533-547, 2006.
82. A. Schulz, A new mixed-integer programming formulation for the maximally diverse grouping problem

with attribute values. Ann. of Oper. Res., 318, 501-530, 2022.
83. A. Schulz, The balanced maximally diverse grouping problem with integer attribute values. J. of Combin. Optim., 45, art. 135, 1–27,, 2023.
84. H.-M. Shieh, M.-D. May, Solving the capacitated clustering problem with genetic algorithms. J. Chin. Inst. Ind. Eng., 18, 1-12, 2001.
85. K. Singh, S. Sundar, A new hybrid genetic algorithm for the maximally diverse grouping problem. Int. J. of Machine Learning and Cybernetics, 10(10), 2901–2940, 2019.
86. W. Tang, Y. Yang, L. Zeng, Y. Zhan, Optimizing MSE for clustering with balanced size constraints. Symmetry, 11, art. 338, 1–16, 2019.
87. D.H. Tran, B. Babaki, D.V. Daele, P. Leyman, P. De Causmaecker, Local search for constrained graph clustering in biological networks. Comp. and Oper. Res., 132, art. 105299, 2021.
88. B. Vallet, B. Soheilian, M. Bredif, Combinatorial clustering and its application to 3D polygonal traffic sigh reconstruction from multiple images. In: ISORS Annals of the Photogrammetry Remote Sensing and Spatial Information Sciences, vol. II-3, pp. 165–172, 2014.
89. R.R. Weitz, M.T. Jelassi, Assigning students to groups: a multi-criteria decision support system approach. Decis. Sci., 23(3), 746–757, 1992.
90. R.R. Weitz, S. Lakshminarayanan, An empirical comparison of heuristic methods for creating maximally diverse groups. J. of the ORS, 25(6), 473–482, 1998.
91. S.T. Wierzchon, M.A. Klopotek, Algorithms of combinatorial cluster analysis. In: Modern Algorithms of Cluster Analysis, Studies in Big Data, vol. 34, Springer, Cham, pp. 67–161, 2018.
92. Y. Xu, P. Guo, Y. Zeng, An iterative neighborhood local search algorithm for capacitated centered clustering problem. IEEE Access, 10, 34497–34510, 2022.
93. Z. Yang, H. Chen, F. Chu, A lagrangian relaxation approach for a large scale new variant of capacitated clustering problem. Comput. Ind. Eng., 61, 430-435, 2011.
94. X. Yang, Z. Cai, T. Jin, Z. Tang, S. Gao, A three-phase search approach with dynamic population size for solving the maximally diverse grouping problem EJOR, 302(3), 925–953, 2022.
95. Q. Zhou, U. Benlic, Q. Wu, J.-K. Hao, Heuristic search to the capacitated clustering problem. EJOR, 273, 464-487, 2019.
96. S. Zhu, D. Wang, T. Li, Data clustering with size constraints. Knowledge-Based Systems, 23(8), 883–889, 2010.

Author work address: Mark Sh. Levin, Inst. for Information Transmission Problem (Kharkevich Institute), Russian Academy of Sciences, 19 Bolshoy Karetny lane, Moscow 127051, Russia

Author home address: Mark Sh. Levin, Sumskoy Proezd 5-1-103, Moscow 117208, Russia

http://www.mslevin.iitp.ru/ email: mslevin@acm.org