What is claimed is:

CLAIMS

1. A computer-implemented method for performing resource allocation, comprising:

using a distributed computing system that includes a number of individual computer-implemented solvers for performing resource allocation of M resources among N users into K pools by solving a knapsack problem (KP) subject to K global constraints and L local constraints:

receiving data representing the K global constraints and the L local constraints, wherein each of the K global constraints restricts a respective maximum per-pool cost of the M resources across two or more users, and each of the L local constraints restricts a per-user selection of the M resources;

transforming the KP into a dual problem using K dual multipliers, each of the K dual multipliers corresponding to a respective one of the K global constraints;

decomposing the dual problem into N sub-problems, each of the N sub-problems corresponding to a respective one of the N users and subject to the L local constraints with respect to (w.r.t.) the respective one of the N users, wherein each of the L constraints corresponds to a different subset of M resources and restricts a respective maximum number C of resources to be selected among the subset of the M resources;

determining initial values of the K dual multipliers, wherein the initial values of the K dual multipliers are computed based on a sampled version of the KP for resource allocation among S users into the K pools subject to K sampled global constraints and the L local constraints, wherein the S users are sampled from the N users, and each of the K sampled global constraints restricts a respective scaled-down maximum total cost of resources across the S users for each of the K pools;

performing two or more iterations in solving the dual problem, wherein in one iteration, for each dual multiplier corresponding to a global constraint corresponding to a pool that restricts a maximum per-pool cost of the M resources across two or more users:

determining an updated dual multiplier for the global constraint corresponding to the pool to be a non-negative threshold such that:

a total cost of resources selected across the N users for the pool corresponding to a dual multiplier that is less than the non-negative threshold exceeds the maximum per-pool cost; and

a total cost of resources selected across the N users for the pool corresponding to a dual multiplier that is not less than the non-negative threshold does not exceed the maximum per-pool cost; and

computing M decision variables of each of the N users corresponding to the updated dual multiplier in solving each of the N sub-problems corresponding to the each of the N users by one of the number of individual computer-implemented solvers, wherein each of the M decision variables indicates whether or not to select a respective one of the M resources by the each of the N users.

2. The method of claim 1, wherein N is in an order of billions or larger, K is smaller than L, L is in an order of billions or larger.

3. The method of any preceding claim, wherein any subset corresponding to a first constraint out of the L constraints has no common resource with any other subset corresponding to a second constraint out of the L constraints, unless the subset corresponding to the first constraint is a subset of the other subset corresponding to the second constraint.

4. The method of any preceding claim, further comprising:

in response to determining that a convergence condition is met,

computing M decision variables of each user of the N users corresponding to the K dual multipliers; and

allocating the M resources among the N users according to the M decision variables of each user of the N users.

5. The method of any preceding claim, wherein solving each of the N sub-problems corresponding to the each of the N users comprises: for a sub-problem corresponding to a user and for each dual multiplier corresponding to the global constraint corresponding to the pool,

computing candidates of the dual multiplier;

sorting the candidates of the dual multiplier in a non-increasing order;

traversing the candidates of the dual multiplier in the non-increasing order:

for each candidate of the dual multiplier,

computing M decision variables of the user corresponding to the candidate of the dual multiplier; and

computing an incremental cost of selected resources based on the M decision variables of the user corresponding to the candidate of the dual multiplier relative to a cost of selected resources based on the M decision variables of the user corresponding to a candidate of the dual multiplier preceding the candidate of the dual multiplier in the non-increasing order; and

wherein a total cost of resources selected across the N users for the pool corresponding to a certain dual multiplier is computed by summing incremental costs of selected resources of the user corresponding to all candidates of the dual multiplier that are larger than or equal to the certain dual multiplier.

6. The method of claim 5, wherein computing candidates of the dual multiplier comprises computing pair-wise intersections points of M linear functions of the dual multiplier, wherein the M linear functions of the dual multiplier represent M cost-adjusted rewards of the M resources given the dual multiplier.

7. The method of claim 5, further comprising computing the non-negative threshold based on the candidates of the dual multiplier sorted in the non-increasing order and a respective total cost of resources selected across the N users for the pool corresponding to one of the candidates of the dual multiplier in the non-increasing order.

8. The method of claim 7, further comprising performing interpolation on the candidates of the dual multiplier sorted in the non-increasing order for the computing the non-negative threshold.

9. The method of claim 5, further comprising:

grouping the candidates of the dual multiplier into a plurality of non-uniform budgets;

computing sums of incremental costs of selected resources of the user corresponding to respective candidates of the dual multiplier in the plurality of non-uniform budgets; and

computing the non-negative threshold based on the sums of incremental costs.

10. A system for performing resource allocation, comprising:

one or more processors; and

one or more computer-readable memories coupled to the one or more processors and having instructions stored thereon that are executable by the one or more processors to perform the method of any of claims 1 to 9.

11. An apparatus for performing resource allocation, the apparatus comprising multiple modules for performing the method of any of claims 1 to 9.