What is claimed is:

CLAIMS

1. A computer-implemented method for allocating M resources subject to L constraints, the method comprising:

receiving data representing L constraints, wherein each of the L constraints corresponds to a subset of M resources and restricts a respective maximum number C of resources to be selected among the subset of the M resources, 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, and wherein the M resources are user traffic that needs to be allocated between different channels;

determining a topological ordering of the L constraints, wherein the topological ordering defines a traversing order of the L constraints;

selecting all the M resources as an initial selection;

removing resources from the initial selection by traversing each constraint in the topological ordering of the L constraints, wherein the removing comprises: for a constraint l of the L constraints corresponding to a subset S_l of the M resources and restricting a maximum number C_l of resources to be selected out of the subset S_l,

determining currently selected resources in the subset S_l; and

unselecting, among the currently selected resources in the subset S_l, any resource that is not one of C_l resources that have top C_l rewards among the currently selected resources in the subset S_l; and

allocating the selected resources after traversing all the L constraints in the topological ordering of the L constraints;

wherein the method is performed for one of N users in performing resource allocation of the M resources among the N users into K pools subject to K global constraints and the L constraints, wherein each of the L constraints restricts a respective maximum number of resources out of the M resources to be selected by a single user, and each of the K global constraints limits resources allocated for each pool across the N users.

2. The method of claim 1, wherein:

the data representing the L constraints comprises L index subsets, each of the L index subsets corresponds to a respective one of the L constraints and comprises indexes of a subset of the M resources corresponding to the respective one of the L constraints; and

the subset S_l comprises indexes of a subset of the M resources corresponding to the constraint l.

3. The method of any preceding claim, wherein the data representing the L constraints comprises data representing each of the L constraints by a node in a directed acyclic graph (DAG) data structure, and an edge of the DAG represents a relationship between two of the L constraints corresponding to two nodes of the edge.

4. The method of any preceding claim, further comprising: obtaining data representing M rewards, each reward corresponding to one of the M resources respectively.

5. The method of claim 4, further comprising:

sorting the M resources in a non-increasing order according to the M rewards corresponding to the M resources before traversing each subset in the L subsets in the topological ordering of the L subsets; and

wherein unselecting, among the currently selected resources in the subset S_l, any resource that is not one of C_l resources that have top C_l rewards among the currently selected resources in the subset S_l comprises unselecting the any resource that is not one of C_l resources that have top C_l rewards among the currently selected resources in the subset S_l according to the non-increasing order of the M resources.

6. The method of claim 4, wherein the M rewards are M cost-adjusted rewards respectively, wherein each of the M cost-adjusted rewards comprises a reward minus a cost if one of the M resources is selected.

7. A system for performing resource allocation of M resources subject to L constraints, 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 6.

8. An apparatus for performing resource allocation of M resources subject to L constraints, the apparatus comprising multiple modules for performing the method of any of claims 1 to 6.