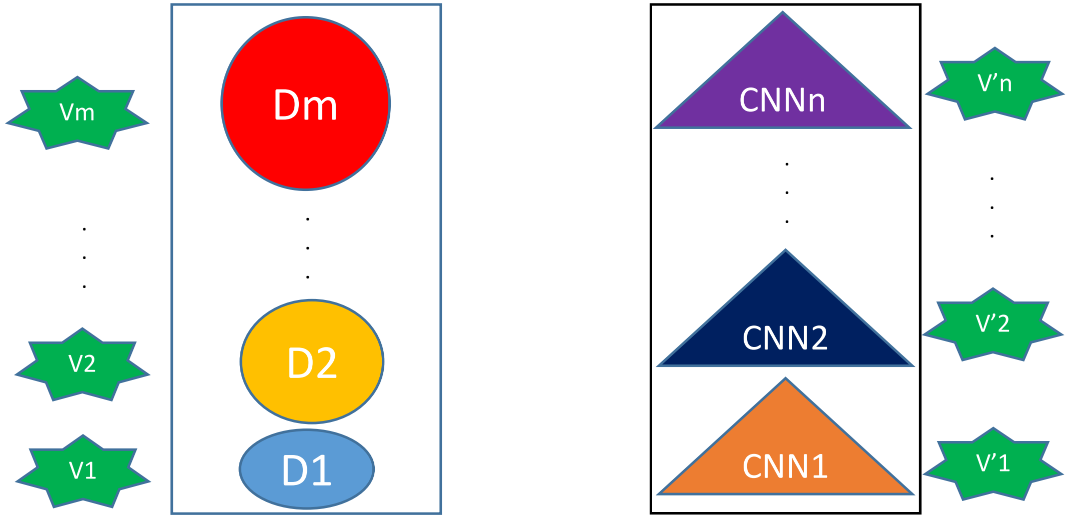
2 Market-based Resource Allocation Problems

Sayed Saghaian

Problem 1: Definition

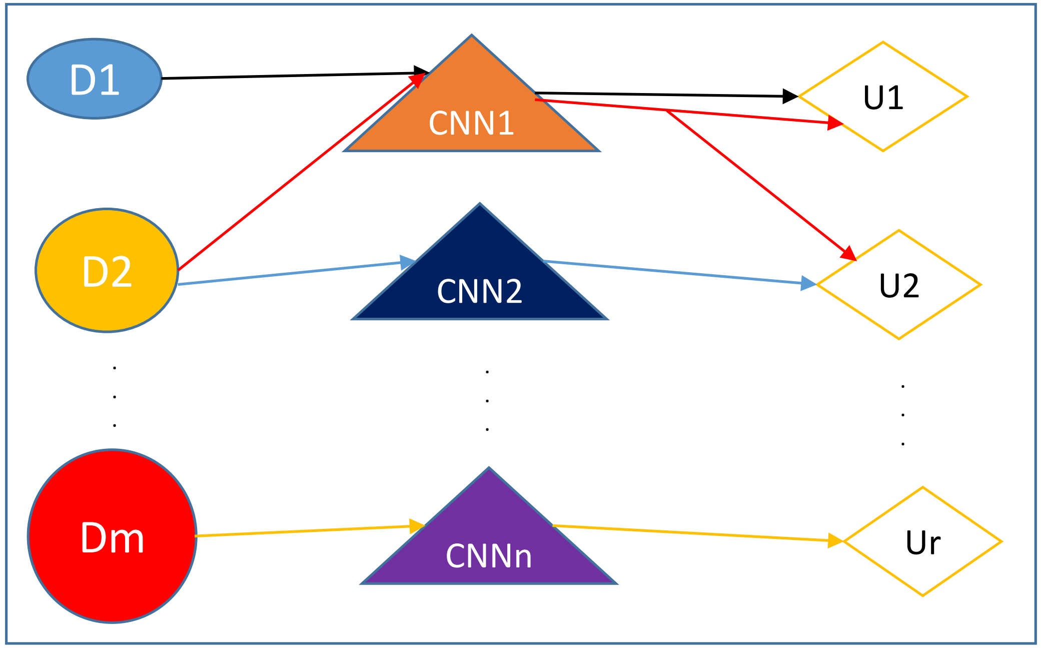
- Bank of datasets and CNN's

$\mathbb{D} = \{D_1, \dots, D_m\}$ $\qquad$ $\mathbb{C} = \{C_1, \dots, C_n\}$



From the "*system*" perspective, each item $j \in \mathbb{D} \cup \mathbb{C}$ has a value V_j (e.g. how *hot* the item is based on previous history)

- Set of requests corresponding (logical) graph:
  - U1 = {CNN1(D1), CNN1(D2)}
  - U2 = {CNN1(D2), CNN2(D2)}
  - Um = {CNNn(Dm)}



Problem 1: Definition

- The physical network consist of nodes that have limited *memory capacities* to store objects as well as limited *power to execute* a CNN operation:
  - Cannot place the entire bank of datasets and CNN's on to the physical network

- Problem:
  - What to Pack → Resource Allocation Problem
  - Where to Pack → Placement Problem
  - Such that total values is maximized while total costs is minimized

Problem1: Assumptions

- D_i's and C_j's are agile (moveable).
- U_r's are anchored to their corresponding physical node.
- C_j's are not decomposable (the whole C_j should be placed on one node).
- From the "*system*" perspective, each item $j \in \mathbb{D} \cup \mathbb{C}$ has a *value* V_j:
  - The value of item j (V_j) is independent of where the other items have been placed.
  - The value of item j (V_j) is independent of the location in the physical network that it resides.
- We assume values and cost have the same unit, otherwise the problem becomes a *Multi Objective Optimization* problem which is more complicated.

Problem1: Assumptions

- From the "*system*" perspective, each item $j \in \mathbb{D} \cup \mathbb{C}$ has a *cost*:
  a. cost of item j itself
  b. cost of item j if the other endpoint item is not placed on the same physical node
- Cost can be think of the communication energy required to:
  - Upload item j (or move in case the item j was already been placed) from its initial location to its optimal location (Thethering cost) ⇒ denoted by: q_{jz}
  - Send the data along the outgoing edges of the logical graph ⇒ denoted by: q_{jzy}
- The cost of C_j is independent of the dataset it will be executed on
  - E.g. the output size of $C_j(D_i)$ is always the same regardless of the D_i's
- The cost of item j (q_{jzy}) does dependent on where its endpoint items are placed
  - Item j is placed on physical node *z* whereas the other endpoint item is placed at node *y*

Problem1: Other Notations

Decision variables x_{jz}:

$$x_{jz} = \begin{cases} 1, \text{if item j is placed at phyisical node z} \\ 0, \text{otherwise} \end{cases}$$

w_{1j}: the required memory to store item j

w_{2j}: the require power to execute item j

c_{1k}: the memory capacity of physical node k

c_{2k}: the power capacity to execute an operation at physical node k

$\mathbb{E}_{j\backslash\mathbb{U}}$: the set of items (not in the set of users) where there is an edge from item j ∈ $\mathbb{D} \cup \mathbb{C}$ to them in the logical graph:

E.g. $\mathbb{E}_{D_2} = \{C_1, C_2\}$

$\mathbb{E}_{j\rightarrow\mathbb{U}}$: the set of users where there is an edge from item j ∈ $\mathbb{C}$ to them in the logical graph:

E.g. $\mathbb{E}_{C_1} = \{U_1, U_2\}$

$\mathbb{P}$: the set of physical nodes

$\mathcal{L}(u) \in \mathbb{P}$: the location of user u in the physical network

Problem 1: Formulation

- This problem is a Quadratic Multiple Multidimensional Knapsack Problem:

$$\max \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}} V_j x_{jz} - \sum_{z,y\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}} \sum_{l\in\mathbb{E}_{j\setminus\mathbb{U}}} q_{jzy} x_{jz} x_{ly} - \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{C}} \sum_{u\in\mathbb{E}_{j\to\mathbb{U}}} q_{jz\mathcal{L}(u)} x_{jz} - \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}} q_{jz} x_{jz}$$

s.t.

$$\sum_{j\in\mathbb{D}\cup\mathbb{C}} w_{ij} x_{jz} \le c_{iz} \qquad \text{for } i = 1,2 \text{ and } \forall z \in \mathbb{P} \text{ (capacity constraints)}$$

$$\sum_{z\in\mathbb{P}} x_{jz} \le 1 \qquad \forall j \in \mathbb{D} \cup \mathbb{C} \text{ (place an item only on one location)}$$

$$x_{jz} \in \{0,1\} \qquad \forall j \in \mathbb{D} \cup \mathbb{C} \text{ and } z \in \mathbb{P}$$

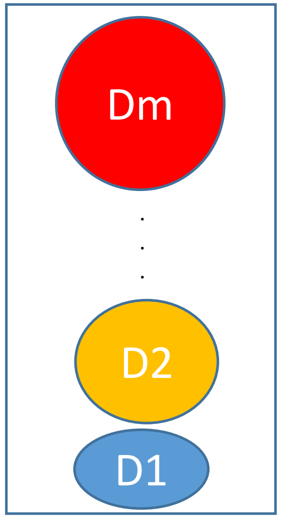
Problem 1: Possible Classes of Solution

1) Tabu Search
2) Genetic Algorithms
3) Simulated Annealing Algorithm
4) Greedy-type Heuristics
5) Relaxation Techniques
6) Exact Solutions (using commercial solvers)

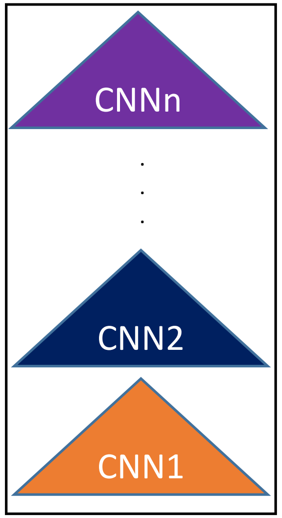
Problem 2: Definition

- Bank of datasets and CNN's

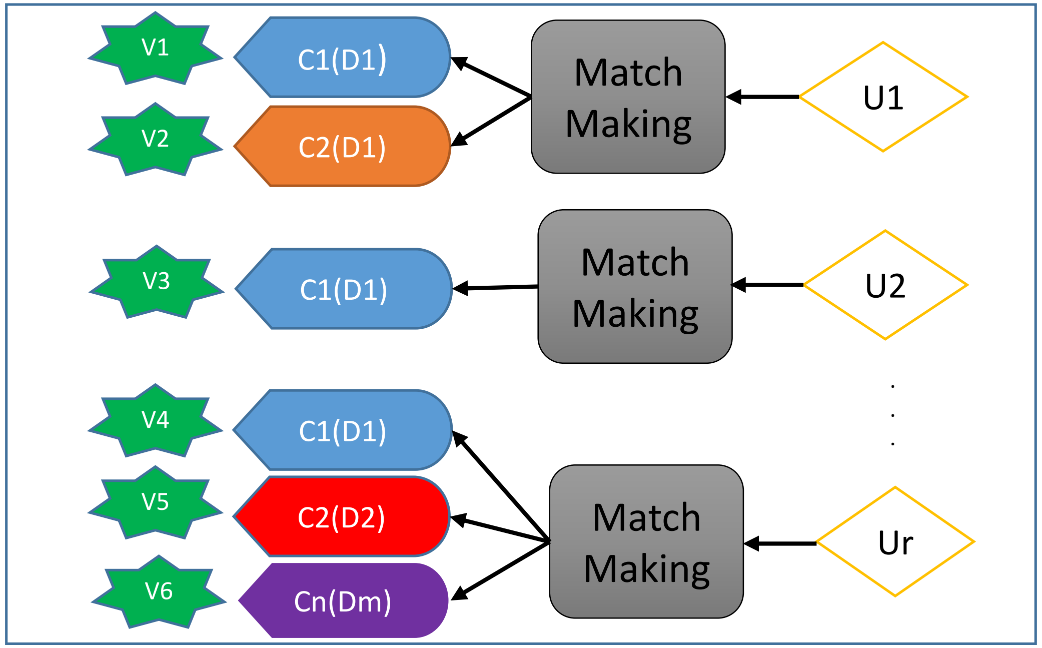
$\mathbb{D} = \{D_1, \dots, D_m\}$



$\mathbb{C} = \{C_1, \dots, C_n\}$



- Incorporating the results of Match Making



- The *bundle* $S = (C_j, D_i)$ corresponding to user k has a value V_{Sk} from this "*user*" perspective
- Cost is defined from the "*system*" perspective

Problem1: **Values** are defined for *each item* and from *system* perspective
Problem2: **Values** are defined on *bundles* and from *users* perspective

Problem 2: Definition

- The physical network consist of nodes that have limited *memory capacities* to store objects as well as limited *power to execute* a CNN operation:
  - Cannot place the entire bank of datasets and CNN's on to the physical network

- The Match Making Algorithm determines a set of bundles that can satisfy a user's request. Each bundle might result in a different value for the *user*.
  - By the bundle $S = (C_j, D_i)$, we refer to the execution of the CNN C_j on the dataset D_i.
  - $S(1) = C_j$ and $S(2) = D_i$ are the constructing elements of the bundle S

- For each user, we should select at most one bundle.
- If a bundle is selected, its constructing elements must be placed (possibly at different nodes of the physical network).
- Placement of both constructing elements of a bundle does not imply this bundle is selected.

- Problem:
  - ➢ Which bundles to select
  - ➢ What objects to Pack ➡ Resource Allocation Problem
  - ➢ Where to Pack ➡ Placement Problem
  - ➢ Such that total values from users is maximized while total costs of the system is minimized

Problem2: Assumptions

- D_i's and C_j's are agile (moveable).
- U_r's are anchored to their corresponding physical node.
- C_j's are not decomposable (the whole C_j should be placed on a one node).
- If *both* constructing elements of the bundle $S = (C_j, D_i)$ corresponding to the user k are placed on the physical network, this user can achieve a value of V_{Sk}.
- If *part* of a bundle corresponding to a user is not placed on the physical network, this user achieves No value from this bundle at all.

Problem2: Other Notations

Decision variable x_{jz}:

$$x_{jz} = \begin{cases} 1, \text{if item j is placed at phyisical node z} \\ 0, \text{otherwise} \end{cases}$$

Decision variable t_{Sk}:

$$t_{Sk} = \begin{cases} 1, \text{if boundle S is selected for user } k \\ 0, \text{otherwise} \end{cases}$$

Auxiliary variable o_S: used to implement the logical OR: $o_S = \bigvee_{k \in \mathbb{U}} t_{Sk}$ (if bundle *S* is selected for at lease one user)

$\mathbb{B}_k$: the set of bundles for user *k* determined by Match Making algorithm

$\mathbb{U}$: the set of users

c_{1k}: the memory capacity of physical node k

c_{2k}: the power capacity to execute an operation at physical node k

Problem 2: Formulation: Cubic Mixed Integer Program

$$\max \sum_{k\in\mathbb{U}} \sum_{S\in\mathbb{B}_k} V_{Sk}\, t_{Sk} - \sum_{S\in\cup\mathbb{B}_k} \sum_{z,y\in\mathbb{P}} q_{S(1)zy}\, o_S\, x_{S(1)z} x_{S(2)y} - \sum_{k\in\mathbb{U}} \sum_{S\in\mathbb{B}_k} \sum_{z\in\mathbb{P}} q_{S(2)z\mathcal{L}(k)}\, t_{Sk}\, x_{S(2)z} - \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}} q_{jz} x_{jz}$$

s.t.

$$\sum_{j\in\mathbb{D}\cup\mathbb{C}} w_{ij} x_{jz} \le c_{iz} \qquad \text{for } i = 1{,}2 \text{ and } \forall z \in \mathbb{P} \text{ (capacity constraints)}$$

$$\sum_{z\in\mathbb{P}} x_{jz} \le 1 \qquad \forall j \in \mathbb{D} \cup \mathbb{C} \text{ (place an item only on one location)}$$

$$\sum_{S\in\mathbb{B}_k} t_{Sk} \le 1 \qquad \forall k \in \mathbb{U} \text{ (select at most one bundle)}$$

$$t_{Sk} \le \sum_{z\in\mathbb{P}} x_{S(1)z} \qquad \forall k \in \mathbb{U} \text{ and } S \in \mathbb{B}_k$$

$$t_{Sk} \le \sum_{z\in\mathbb{P}} x_{S(2)z} \qquad \text{(if a bundle is selected, its constructing elemetns must be placed)}$$

$$o_S \ge t_{Sk} \qquad \forall S \in \bigcup_{k\in\mathbb{U}} B_k \text{ and } k \in \mathbb{U}$$

$$o_S \le \sum_{k\in\mathbb{U}} t_{Sk} \qquad \text{(logcal OR)}$$

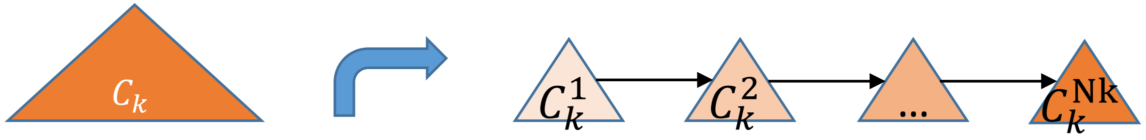
$$x_{jz} \in \{0,1\} \qquad \forall j \in \mathbb{D} \cup \mathbb{C} \text{ and } z \in \mathbb{P}$$
$$t_{Sk} \in \{0,1\} \qquad \forall k \in \mathbb{U} \text{ and } S \in \mathbb{B}_k$$
$$0 \le o_S \le 1 \qquad \forall S \in \bigcup \mathbb{B}_k$$

Problem 2: Possible Classes of Solution

1) Combinatorial Auction Design
2) Tabu Search
3) Genetic Algorithms
4) Simulated Annealing Algorithm
5) Greedy-type Heuristics
6) Relaxation Techniques
7) Exact Solutions (using commercial solvers)

Allowing Decomposition of CNNs

- In this extension, we allow C_k's to be *decomposable*:
  - the constructing elements of a C_k can be placed on different physical nodes
- A C_k is constructed of N_k elements (layers): $C_k = (C_k^1, C_k^2, \dots, C_k^{N_k})$



Let $\mathbb{C}^A = \bigcup_{k=1}^{n} \bigcup_{h=1}^{N_k} C_k^h$ denote the set of constructing elements of all C_k's

a) Partial Placement

- Placement of a layer provides additional value
- $V_{C_k^h}$: Additional value that the element C_k^h provides: $V_{C_k} = \sum_{h=1}^{N_k} V_{C_k^h}$
- $x_{C_k^h z} = 1$: Place C_k^h at the physical node z
- If a layer is placed, all the previous layers must be placed too

b) All or Nothing Placement:

- If all parts of a C_k are placed, the value V_{C_k} for this C_k is achieved
- If any part of a C_k is not placed, No value for this C_k is achieved
- $V_{C_k^1} = V_{C_k^2} = \cdots = V_{C_k^{N_k - 1}} = 0$
- $V_{C_k^{N_k}} = V_{C_k}$

Problem 1: a) Partial Placement

$$\max \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}^A} V_j x_{jz}$$
$$- \sum_{z,y\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}^A} \sum_{l\in\mathbb{E}_{j\setminus\mathbb{U}}} q_{jzy} x_{jz} x_{ly} - \left[\sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{C}} \sum_{u\in\mathbb{E}_{j\to\mathbb{U}}} q_{j^{N_j} z \mathcal{L}(u)} x_{j^{N_j} z} + \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{C}} \sum_{h=1}^{N_j-1} \sum_{u\in\mathbb{E}_{j\to\mathbb{U}}} q_{j^h z \mathcal{L}(u)} x_{j^h z} \left(1 - \sum_{y\in\mathbb{P}} x_{j^{h+1} y}\right) \right]$$
$$- \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}^A} q_{jz} x_{jz}$$

s.t.

$$\sum_{j\in\mathbb{D}\cup\mathbb{C}^A} w_{ij} x_{jz} \le c_{iz} \qquad \text{for } i = 1,2 \text{ and } \forall z \in \mathbb{P} \ \text{(capacity constraints)}$$

$$\sum_{z\in\mathbb{P}} x_{jz} \le 1 \qquad \forall j \in \mathbb{D} \cup \mathbb{C}^A \ \text{(place an item only on one location)}$$

$$\sum_{z\in\mathbb{P}} x_{j^{N_j} z} \le \cdots \le \sum_{z\in\mathbb{P}} x_{j^2 z} \le \sum_{z\in\mathbb{P}} x_{j^1 z} \qquad \forall j \in \mathbb{C} \ \text{(If a layer is placed, all the previous layers must be placed too)}$$

$$x_{jz} \in \{0,1\} \qquad \forall j \in \mathbb{D} \cup \mathbb{C}^A \text{ and } z \in \mathbb{P}$$

Problem 1: b) All or Nothing Placement

$$\max \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}^A} V_j x_{jz} - \sum_{z,y\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}^A} \sum_{l\in\mathbb{E}_{j\setminus\mathbb{U}}} q_{jzy} x_{jz} x_{ly} - \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{C}} \sum_{u\in\mathbb{E}_{j\to\mathbb{U}}} q_{j^{Nj}z\mathcal{L}(u)} x_{j^{Nj}z} - \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}^A} q_{jz} x_{jz}$$

s.t.

$$\sum_{j\in\mathbb{D}\cup\mathbb{C}^A} w_{ij} x_{jz} \le c_{iz} \qquad \text{for } i = 1{,}2 \text{ and } \forall z \in \mathbb{P} \text{ (capacity constraints)}$$

$$\sum_{z\in\mathbb{P}} x_{jz} \le 1 \qquad \forall j \in \mathbb{D} \cup \mathbb{C}^A \text{ (place an item only on one location)}$$

$$\sum_{z\in\mathbb{P}} x_{j^{Nj}z} \le \cdots \le \sum_{z\in\mathbb{P}} x_{j^2z} \le \sum_{z\in\mathbb{P}} x_{j^1z} \qquad \forall j \in \mathbb{C} \text{ (If a layer is placed, all the previous layers must be placed too)}$$

$$x_{jz} \in \{0,1\} \qquad \forall j \in \mathbb{D} \cup \mathbb{C}^A \text{ and } z \in \mathbb{P}$$

Problem 2: a) Partial Placement

$$\max \sum_{z\in\mathbb{P}} \sum_{k\in\mathbb{U}} \sum_{S\in\mathbb{B}_k} \sum_{h=1}^{N_{S(2)}} V_{S(2)^h k}\, t_{Sk} x_{S(2)^h z} - \left[\sum_{S\in\cup\mathbb{B}_k} \sum_{z,y\in\mathbb{P}} q_{S(1)zy}\, o_S\, x_{S(1)z} x_{S(2)^1 y} + \sum_{S\in\cup\mathbb{B}_k} \sum_{z,y\in\mathbb{P}} \sum_{h=1}^{N_{S(2)}-1} q_{S(2)^h zy}\, o_S\, x_{S(2)^h z} x_{S(2)^{h+1} y} \right]$$
$$- \left[\sum_{k\in\mathbb{U}} \sum_{S\in\mathbb{B}_k} \sum_{z\in\mathbb{P}} q_{S(2)^{N_{S(2)}} z\mathcal{L}(k)}\, t_{Sk}\, x_{S(2)^{N_{S(2)}} z} + \sum_{k\in\mathbb{U}} \sum_{S\in\mathbb{B}_k} \sum_{z\in\mathbb{P}} \sum_{h=1}^{N_{S(2)}-1} q_{S(2)^h z\mathcal{L}(k)}\, t_{Sk}\, x_{S(2)^h z} \left(1 - \sum_{y\in\mathbb{P}} x_{S(2)^{h+1} y} \right) \right] - \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}^A} q_{jz} x_{jz}$$

s.t.

$$\sum_{j\in\mathbb{D}\cup\mathbb{C}^A} w_{ij} x_{jz} \le c_{iz} \qquad \text{for } i = 1,2 \text{ and } \forall z \in \mathbb{P} \text{ (capacity constraints)}$$

$$\sum_{z\in\mathbb{P}} x_{jz} \le 1 \qquad \forall j \in \mathbb{D} \cup \mathbb{C}^A \text{ (place an item only on one location)}$$

$$\sum_{S\in\mathbb{B}_k} t_{Sk} \le 1 \qquad \forall k \in \mathbb{U} \text{ (select at most one bundle)}$$

$$t_{Sk} \le \sum_{z\in\mathbb{P}} x_{S(1)z} \qquad \forall k \in \mathbb{U} \text{ and } S \in \mathbb{B}_k$$

$$t_{Sk} \le \sum_{z\in\mathbb{P}} x_{S(2)^1 z} \qquad \text{(if a bundle is selected, the constructing data \& first layer of the CNN must be placed)}$$

$$o_S \ge t_{Sk} \qquad \forall S \in \bigcup_{k\in\mathbb{U}} B_k \text{ and } k \in \mathbb{U}$$

$$o_S \le \sum_{k\in\mathbb{U}} t_{Sk} \qquad \text{(logcal OR)}$$

$$\sum_{z\in\mathbb{P}} x_{j^{N_j} z} \le \cdots \le \sum_{z\in\mathbb{P}} x_{j^2 z} \le \sum_{z\in\mathbb{P}} x_{j^1 z} \qquad \forall j \in \mathbb{C} \text{ (If a layer is placed, all the previous layers must be placed too)}$$

$$x_{jz} \in \{0,1\} \qquad \forall j \in \mathbb{D} \cup \mathbb{C}^A \text{ and } z \in \mathbb{P}$$
$$t_{Sk} \in \{0,1\} \quad \forall k \in \mathbb{U} \text{ and } S \in \mathbb{B}_k$$
$$0 \le o_S \le 1 \quad \forall S \in \bigcup \mathbb{B}_k$$

Problem 2: b) All or Nothing Placement

$$\max \sum_{z\in\mathbb{P}} \sum_{k\in\mathbb{U}} \sum_{S\in\mathbb{B}_k} V_{S(2)k}\, t_{Sk} x_{S(2)^{N_{S(2)}}z} - \left[\sum_{S\in\cup\mathbb{B}_k} \sum_{z,y\in\mathbb{P}} q_{S(1)zy}\, o_S\, x_{S(1)z} x_{S(2)^1 y} + \sum_{S\in\cup\mathbb{B}_k} \sum_{z,y\in\mathbb{P}} \sum_{h=1}^{N_{S(2)}-1} q_{S(2)^h zy}\, o_S\, x_{S(2)^h z} x_{S(2)^{h+1} y} \right]$$
$$- \sum_{k\in\mathbb{U}} \sum_{S\in\mathbb{B}_k} \sum_{z\in\mathbb{P}} q_{S(2)^{N_{S(2)}} z \mathcal{L}(k)}\, t_{Sk}\, x_{S(2)^{N_{S(2)}}z} - \sum_{z\in\mathbb{P}} \sum_{j\in\mathbb{D}\cup\mathbb{C}^A} q_{jz} x_{jz}$$

s.t.

$$\sum_{j\in\mathbb{D}\cup\mathbb{C}^A} w_{ij} x_{jz} \le c_{iz} \qquad \text{for } i = 1,2 \text{ and } \forall z \in \mathbb{P} \text{ (capacity constraints)}$$

$$\sum_{z\in\mathbb{P}} x_{jz} \le 1 \qquad \forall j \in \mathbb{D} \cup \mathbb{C}^A \text{ (place an item only on one location)}$$

$$\sum_{S\in\mathbb{B}_k} t_{Sk} \le 1 \qquad \forall k \in \mathbb{U} \text{ (select at most one bundle)}$$

$$t_{Sk} \le \sum_{z\in\mathbb{P}} x_{S(1)z} \qquad \forall k \in \mathbb{U} \text{ and } S \in \mathbb{B}_k$$
$$t_{Sk} \le \sum_{z\in\mathbb{P}} x_{S(2)^1 z} \quad \text{(if a bundle is selected, the constructing data \& first layer of the CNN must be placed)}$$
$$o_S \ge t_{Sk} \qquad \forall S \in \bigcup_{k\in\mathbb{U}} B_k \text{ and } k \in \mathbb{U}$$
$$o_S \le \sum_{k\in\mathbb{U}} t_{Sk} \quad \text{(logcal OR)}$$
$$\sum_{z\in\mathbb{P}} x_{j^{N_j}z} \le \cdots \le \sum_{z\in\mathbb{P}} x_{j^2 z} \le \sum_{z\in\mathbb{P}} x_{j^1 z} \qquad \forall j \in \mathbb{C} \text{ (If a layer is placed, all the previous layers must be placed too)}$$
$$x_{jz} \in \{0,1\} \qquad \forall j \in \mathbb{D} \cup \mathbb{C}^A \text{ and } z \in \mathbb{P}$$
$$t_{Sk} \in \{0,1\} \quad \forall k \in \mathbb{U} \text{ and } S \in \mathbb{B}_k$$
$$0 \le o_S \le 1 \quad \forall S \in \bigcup \mathbb{B}_k$$