

---

# MSBA7003 Decision Analytics



ZHANG, Wei  
Associate Professor  
HKU Business School

## 07 Mathematical Programming III

# Improving Delivery Networks at JD.com

- Introduction
- Model for Improving Delivery Networks
- Practical Considerations
- Model Results
- Implementation
- Impact and Future Plans



# Company Background

- JD.com: one of top two B2C e-commerce companies in China.
- Known for advanced logistics and high-quality deliver-to-door services.
- Owns the largest amount of logistics infrastructure among all Chinese e-commerce companies.
- By the end of 2018, JD.com had 16 “Asia No.1” logistics parks, 550 large warehouses, and 7,000 delivery stations, serving 99% of China’s population.
- Around 2020, JD.com had undertaken several market expansions and experienced a general increase in demand.



# Company Background

- Revenue growth from 2018 to 2020:

	For the Year Ended December 31,						
	2018		2019		2020		
	RMB	%	RMB	%	RMB	US \$	%
	(in millions, except for percentages)						
Electronics and home appliances revenues	280,059	60.6	328,703	57.0	400,927	61,445	53.8
General merchandise revenues	136,050	29.5	182,031	31.5	250,952	38,460	33.6
Net product revenues	416,109	90.1	510,734	88.5	651,879	99,905	87.4
Marketplace and marketing revenues	33,532	7.2	42,680	7.4	53,473	8,195	7.2
Logistics and other service revenues	12,379	2.7	23,474	4.1	40,450	6,199	5.4
Net service revenues	45,911	9.9	66,154	11.5	93,923	14,394	12.6
Total net revenues	462,020	100.0	576,888	100.0	745,802	114,299	100.0



# Motivation



- As a result of its expansions, JD.com realized that its existing delivery system will not be able to keep up with customer service targets in the future.
- The mismatch between supply and demand is especially large during “6.18” and “double 11”.

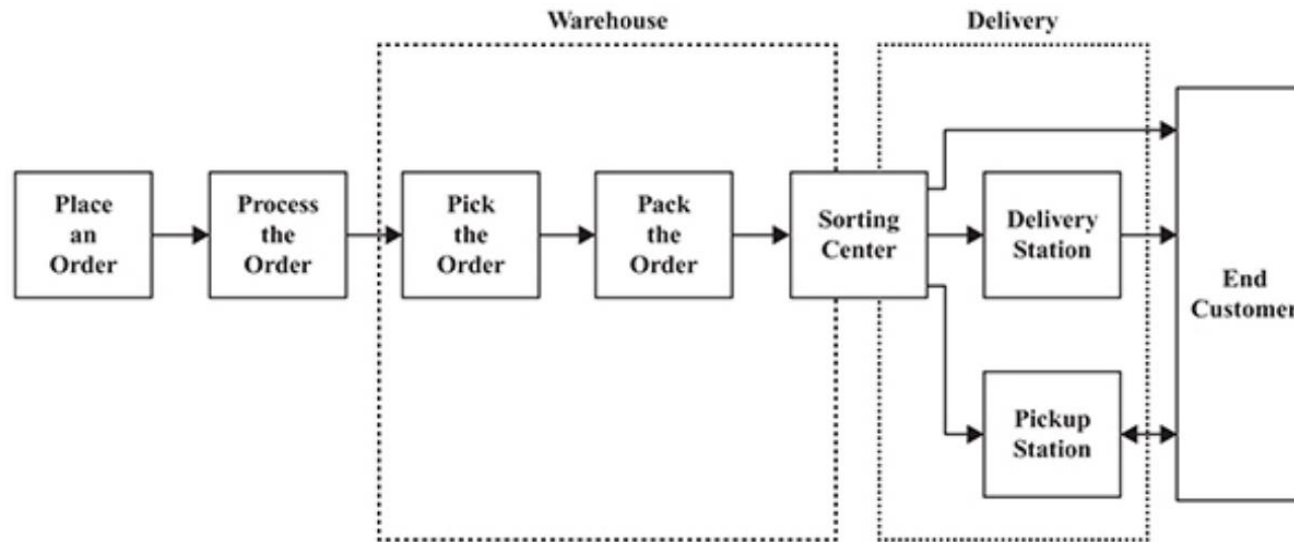
# Motivation

- JD Logistics became an independent and open logistics platform in early 2017.



# Motivation

- Around 2019~20, A change in the structure of its delivery system was necessary to maintain customer service and lower operational costs.

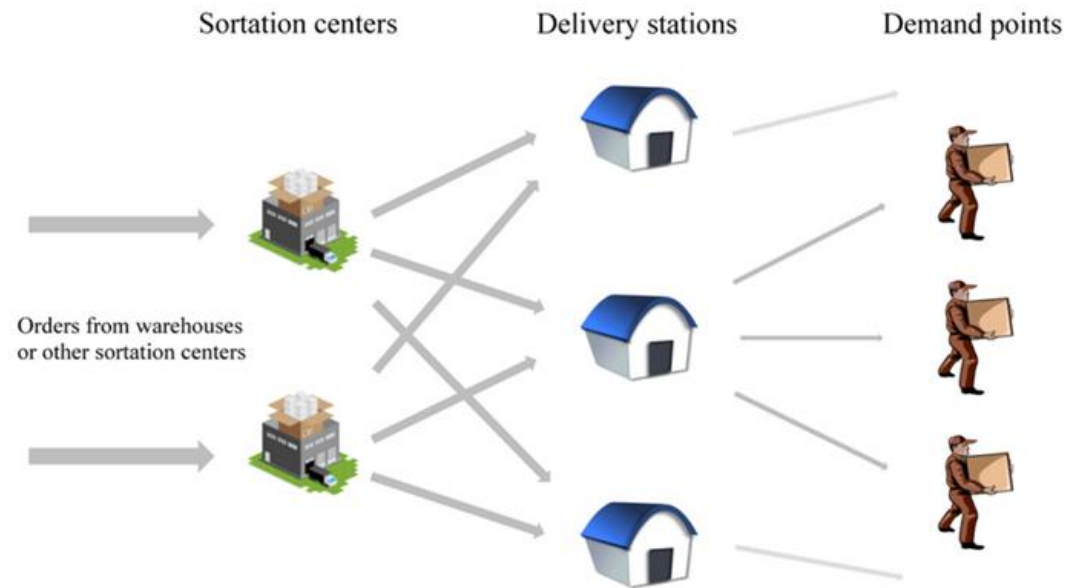


- The key was adjusting citywide delivery station (DS) locations.
  - DSs form the last layer of the logistics system and have pronounced impact.
  - DSs are leased on an annual basis and thus changes can be made easily.

# Motivation

- A citywide distribution system consists of three layers of vertices: sortation centers (SCs), delivery stations (DSs), and blocks (represented by demand points), as well as transportation arcs between them.

An SC receives packages from warehouses or other SCs, sorts them by their destinations, and sends them to corresponding DSs.



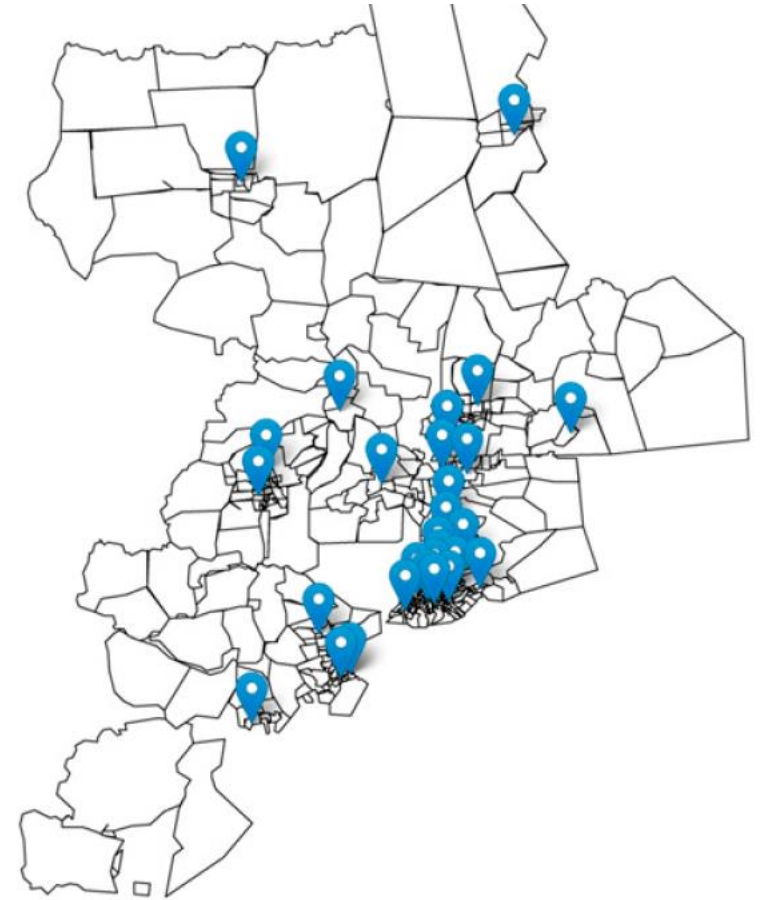
JD.com geographically partitions a city into blocks, and one delivery man typically covers a single block. Each DS delivers packages to its surrounding blocks.

A DS receives packages from SCs and eventually delivers them to end customers. DS staff consist of a station manager and several delivery men.



# Objective

- Traditional manual planning is time-consuming and error-prone.
- Develop a tool to automatically generate annual plans for Delivery Station (DS) location monthly adjustments.
- The goal is to minimize total yearly operational costs.
- Incorporate data preparation and results visualization.





# Model Parameters

- Imported based on practical circumstances
  - lease durations
  - the maximum number of location and assignment changes
  - predicted delivery demand
- Estimated based on existing data of JD.com
  - DS capacity as a linear function of size
  - operational cost as a linear function of size
- Obtained external sources (like map services)
  - Delivery distance between blocks
  - Connectivity between blocks



# A Basic Model

- Indices:

- $j, k$ : index of a block;  $J$  is the entire set.
- $t$ : index of a period (month);  $T = \{1, 2, \dots, 12\}$  is the entire set of periods.

- Parameters:

- $d_{jk}$ : fixed delivery cost per package between blocks  $j$  and  $k$ .
- $q_{tj}$ : predicted number of packages to be delivered to block  $j$  in month  $t$ .
- $c_{tj}$ : package processing capacity of DS in block  $j$  in month  $t$ .
- $z_{tj} = 2000 + 1.2 \times c_{tj}$ : the fixed cost of operating a DS in block  $j$  in month  $t$ .

- Decision Variables

- $X_{tj} = 1$  if a DS exists in block  $j$  in month  $t$  and  $= 0$  otherwise.
- $Y_{tjk} = 1$  if block  $k$  is assigned to the DS in block  $j$ .

# A Basic Model

- Assume that the DS locations can be changed every month:

$$\min \sum_{t \in T} \sum_{j \in J} z_{tj} \cdot X_{tj} + \sum_{t \in T} \sum_{j \in J} \sum_{k \in J} q_{tk} \cdot d_{jk} \cdot Y_{tjk}$$

- s.t.

$$\begin{aligned} \sum_{j \in J} Y_{tjk} &= 1 & \forall t \in T; \forall k \in J & \quad \text{(Every block must be assigned to one DS)} \\ \sum_{k \in J} Y_{tjk} \cdot q_{tk} &\leq c_{tj} \cdot X_{tj} & \forall t \in T; \forall j \in J & \quad \text{(DS capacity constraint)} \\ X_{tj}, Y_{tjk} &\in \{0,1\} & \forall t \in T; \forall j, k \in J & \end{aligned}$$



# A Basic Model

A company is optimizing its logistic network in a city. The city is divided into 100 blocks. The delivery cost per package between any two blocks is given in the attached file "Session\_07\_JD\_Block\_Delivery\_Cost.csv". The number of packages to be delivered to each block in each of the next twelve months is predicted and listed in the attached file "Session\_07\_JD\_Block\_Demand.csv". The company needs to decide in each block whether a delivery station should be located there in each of the next twelve months. If a delivery station is located in a block in a month, it can serve the nearby blocks by fulfilling the package deliver demand in that month. Hence, the company needs to decide for each block which delivery station will fulfill its delivery demand. The package processing and delivery capacity of each delivery station is limited. In the attached file "Session\_07\_JD\_DS\_Capacity.csv", you can find the capacity of a delivery station if it is located in a specific block. There is a fixed cost of operating a delivery station in a block in each month, which equals 2000 plus 1.2 times the capacity in that month. The company now needs an annual plan of delivery station location and coverage to minimize the total operational cost. Please write a Python script to perform the following tasks. First, read the data files and prepare the data in appropriate forms. Maybe consider turning the delivery cost data into a matrix format. Second, formulate the problem as a mixed integer program with PuLP package. Note that there are two types binary decision variables: one is for delivery station location, and the other is for assigning each block to a delivery station. Third, solve the problem and print the total annual cost.



Session\_07\_JD\_Block\_Delivery\_Cost.csv  
CSV · 185 KB



Session\_07\_JD\_Block\_Demand.csv  
CSV · 5 KB



GPT-4o Poe

To solve this problem, you can follow these steps in Python using the PuLP package. This script will read the data, formulate the problem as a mixed integer program, and solve it to minimize the total operational cost.

python

Copy

```
import pandas as pd
import numpy as np
import pulp

# Load data
delivery_cost_df = pd.read_csv('Session_07_JD_Block_Delivery_Cost.csv')
demand_df = pd.read_csv('Session_07_JD_Block_Demand.csv') # Assuming correct format
capacity_df = pd.read_csv('Session_07_JD_DS_Capacity.csv') # Assuming correct format
```

...

```
# Each block's demand must be fulfilled
for j in range(num_blocks):
    for m in range(months):
        problem += pulp.lpSum(y[(i, j, m)] for i in range(num_blocks)) == 1

# Assignments only to open stations
for i in range(num_blocks):
    for j in range(num_blocks):
        for m in range(months):
            problem += y[(i, j, m)] <= x[(i, m)]

# Capacity constraints
for i in range(num_blocks):
    for m in range(months):
        problem += pulp.lpSum(demand_df.iloc[j, m+1] * y[(i, j, m)] for j in range(num_blocks)) <= capacity_df.iloc[i, m+1] * x[(i, m)]

# Solve the problem
problem.solve()

# Print total annual cost
print("Total Annual Cost:", pulp.value(problem.objective))
```

redundant

Total Annual Cost: 5459469.649400254



# Practical Consideration 1

- Changing the location of a DS means either opening a new station or moving an existing station to a different location, which requires huge (non-monetary) efforts and an expiring contract.
  - Upper limit on DS location changes and block assignment changes per month.
    - No changes can be made in the peak months (e.g., 6.18 and 11.11).
    - The operations team will be shorthanded in holiday season.
    - The operations team has limited energy during a regular month.
  - Monotonicity in total number of DSs.
    - to avoid fluctuations due to seasonal demand shifts
    - In cities with surplus DSs or low DS use rates, the number decreases over months
    - In cities with insufficient DSs, the number ramps up over months
  - Time window for DS location change is open only when the contract expires.
    - DS location change at a block can happen only once a year

# Practical Consideration 1

- New decision variables

- $U_{tj} = 1$  if a DS is added or relocated to block  $j$  in month  $t$  and  $= 0$  otherwise.

- New parameters

- $n_t$ : limits on the number of additions or location changes of all delivery stations during month  $t$ .
- $A_{tj}$ : ( $= 1$  or  $0$ ) whether adding a DS or a change in DS location is allowed in block  $j$  during month  $t$ .
- $X_{0j}$ : ( $= 1$  or  $0$ ) the initial DS locations

- New constraints

- $\sum_{j \in J} U_{tj} \leq n_t \quad \forall t \in T$  (DS addition or relocation cannot exceed  $n_t$  in month  $t$ )
- $X_{tj} - X_{t-1,j} \leq U_{tj} \quad \forall j \in J; \forall t \in T$  (adding a DS  $= 1$  if a DS is added)
- $X_{t-1,j} - X_{tj} \leq A_{tj} \quad \forall j \in J; \forall t \in T$  (removing a DS in block  $j$  is allowed in month  $t$ )
- $U_{tj} \leq A_{tj} \quad \forall j \in J; \forall t \in T$  (adding a DS in block  $j$  is allowed in month  $t$ )
- $\sum_{j \in J} X_{tj} \leq \sum_{j \in J} X_{t-1,j} \quad \forall t \in T$  (monotonicity constraint; assuming surplus DS in the city)
- $U_{tj} \in \{0,1\} \quad \forall t \in T; \forall j, k \in J$

# Practical Consideration 1

- A small example with 3 blocks and 4 months
  - (Assume  $z_{tj} = 0$ ,  $c_{tj} = 100$ ,  $d_{jk} = 2$  for  $k \neq j$ , and  $d_{jj} = 1$ .)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Package Delivery Demand							Month 1 assignment									Distance			
2	t \ j	1	2	3				to \ from	1	2	3			Demand	Capacity		1	1	2	3
3	1	55	20	50				1	1	0	0			55	100		1	1	2	2
4	2	40	30	60				2	0	0	0			0	0		2	2	1	2
5	3	35	40	40				3	0	1	1			70	100		3	2	2	1
6	4	30	60	20					1	1	1	=	1				Month			
7																	1	1	2	1
8	DS location							Month 2 assignment									2	1	2	1
9	j \ t	0	1	2	3	4		to \ from	1	2	3			Demand	Capacity		3	1	2	1
10	1	1	1	1	1	1		1	1	1	0			70	100		4	1	2	1
11	2	1	0	0	0	0		2	0	0	0			0	0		Total Cost			
12	3	0	1	1	1	1		3	0	0	1			60	100		630			
13	Existing	2	2	2	2	2			1	1	1	=	1							
14		>=	>=	>=	>=															
15								Month 3 assignment												
16	Location Change							to \ from	1	2	3			Demand	Capacity					
17	X(t) - X(t-1)	1	0	0	0	0		1	1	1	0			75	100					
18		2	-1	0	0	0		2	0	0	0			0	0					
19		3	1	0	0	0		3	0	0	1			40	100					
20	X(t-1) - X(t)	1	0	0	0	0			1	1	1	=	1							
21		2	1	0	0	0														
22		3	-1	0	0	0		Month 4 assignment												
23	Planned	1	0	0	0	0		to \ from	1	2	3			Demand	Capacity					
24		2	0	0	0	0		1	1	0	0			30	100					
25		3	1	0	0	0		2	0	0	0			0	0					
26	Availability	1	0	0	1	0		3	0	1	1			80	100					
27		2	1	0	0	0			1	1	1	=	1							
28		3	1	1	1	1														
29	# of changes		1	0	0	0														
30	<= Quota		2	2	2	2														



# Practical Consideration 2

- Upper limit on block assignment changes per month.
  - A block assignment change from DS1 to DS2 requires the delivery man responsible for the block has to change his personnel organization from DS1 to DS2, or a delivery man in DS2 has to take over this job. To avoid frequent reallocation of human resources.
- New decision variables
  - $W_{tk} = 1$  if the DS assignment of block  $k$  is changed in month  $t$  and  $= 0$  otherwise.
- New parameters
  - $n$ : limit on the number of assignment changes for each block over a year.
  - $Y_{0jk}$ : the initial assignment
- New constraints
  - $\sum_{t \in T} W_{tk} \leq n \quad \forall k \in J$  (DS assignment for each block cannot exceed  $n$  times per year)
  - $Y_{tjk} - Y_{t-1,jk} \leq W_{tk} \quad \forall j, k \in J; \forall t \in T$  (DS assignment change = 1 if assignment changes)
  - $W_{tk} \in \{0,1\} \quad \forall t \in T; \forall j, k \in J$



## Practical Consideration 3

- JD can adjust the size of a DS and thus the capacity and cost.
- New decision variables
  - $S_{tj} \geq 0$ : the size of DS at block  $j$  in month  $t$
  - $Z_{tj} \geq 0$ : the operational cost of DS at block  $j$  in month  $t$
- New parameters
  - $\alpha^{ca}$ : the constant term in the size-capacity linear function
  - $\beta^{ca}$ : the coefficient in the size-capacity linear function
  - $\alpha^{co}$ : the constant term in the size-cost linear function
  - $\beta^{co}$ : the coefficient in the size-cost linear function
  - $S_{0j}$ : the initial size of DS at block  $j$
  - $s^l, s^u$ : the lower and upper bound of DS size

# Practical Consideration 3

- New constraints

- $Z_{tj} \geq \alpha^{co} + \beta^{co} \cdot S_{tj} - (\alpha^{co} + \beta^{co} \cdot s^u) \cdot (1 - X_{tj}) \quad \forall t \in T; \forall j \in J$
- $Z_{tj} \geq (\alpha^{co} + \beta^{co} \cdot S_{tj}) \cdot X_{tj} ??$  (This is equivalent but not allowed)
- $S_{tj} \geq S_{t-1,j} + s^l \cdot U_{tj} \quad \forall t \in T; \forall j \in J$  (DS size can be chosen only when it is new)
- $S_{tj} \leq S_{t-1,j} + s^u \cdot U_{tj} \quad \forall t \in T; \forall j \in J$  (DS size can be chosen only when it is new)

(DS operational cost depends on both  $X_{tj}$  and  $S_{tj}$  in a linear way)

- Revised constraints

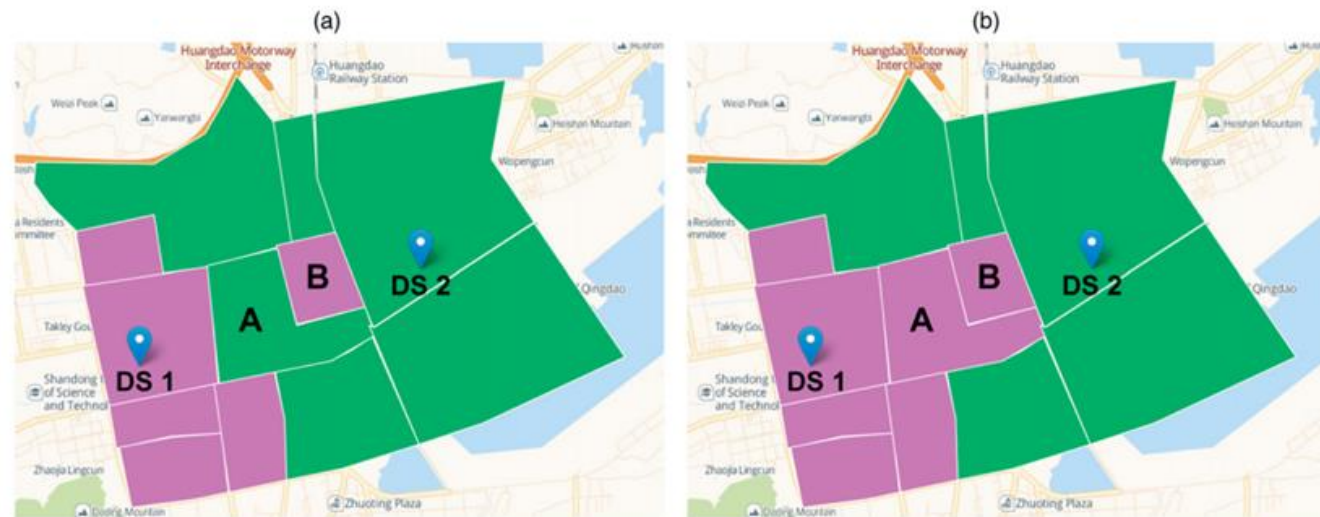
- $\sum_{k \in J} Y_{tjk} \cdot q_{tk} \leq \alpha^{ca} \cdot X_{tj} + \beta^{ca} \cdot S_{tj} \quad \forall t \in T; \forall j \in J$  (DS capacity constraint)

- Revised objective

- $\min \quad \sum_{t \in T} \sum_{j \in J} Z_{tj} + \sum_{t \in T} \sum_{j \in J} \sum_{k \in J} q_{tk} \cdot d_{jk} \cdot Y_{tjk}$

# Practical Consideration 4

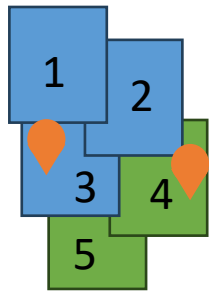
- Block adjacency: two blocks are “adjacent” if they share at least one side or vertex.
- Path of blocks: a sequence of blocks in which successive pairs of blocks are adjacent.
- Block connectivity: there exists a path between them.
- All blocks assigned to the same DS must be connected.





# Practical Consideration 4

- A small example



The following constraints are imposed between any two blocks. If block 1 and 2 are assigned to the same DS, there should be a path between them. For the block on the path, the number of entries must equal the number of exits.

	A	B	C	D	E	F	G	H	I	J
1	Adjacency Matrix									
2		1	2	3	4	5				
3	1	1	1	1	0	0				
4	2	1	1	1	1	0				
5	3	1	1	1	1	1				
6	4	0	1	1	1	1				
7	5	0	0	1	1	1				
8										
9	Connectivity Constraints between 1 and 2									
10	from \ to	1	2	3	4	5	# of Exit			
11	1	0	0	1	0	0	1	=sumproduct(B3:F3,B11:F11)		
12	2	0	0	0	0	0	0			
13	3	0	1	0	0	0	1			
14	4	0	0	0	0	0	0			
15	5	0	0	0	0	0	0			
16	# of Entry	0	1	1	0	0	=sumproduct(F3:F7,F11:F15)			
17	# of Exit	1	0	1	0	0				
18	Start	Y_t31	0	0	0	0				
19	End	0	Y_t32	0	0	0				
20	Balance	=0	=0	=0	=0	=0	= (# of Entry + Start) - (# of Exit + End)			



# Challenges

- A huge number of variables and constraints in the optimization model.
  - Example: Qingdao has 545 candidate DS locations and equal number of demand points.
  - An integer programming model with variables and constraints in the millions.
  - For Beijing and Shanghai, the size is even bigger.
- It is difficult to solve the model by using existing solvers or methodologies.
  - A specialized solution method is developed.



# Solution Approach

- Blocks are more likely to be assigned to a nearby DS
- $\tau\%$ -neighborhood-restricted ( $\tau\%$ -NR) problem to find near-optimal solutions
- $\tau = 100$  for the original problem
- Choose  $\tau$  to balance between optimality and computation efficiency
- $\tau$  values are tried in an increasing order and problems are solved in a nested way
  - Consider 5%-, 6%-, and 7%-NR problems
  - The 5%-NR problem is solved to optimality first
  - The optimal solution is used as a warm start (or the default setting) for the 6%-NR problem
  - Repeat the above procedure to solve the 7%-NR problem

# Model Results (Case study: Annual planning for Qingdao)

- Significant reduction in computation time and close-to-optimal solution with the 10%-NR problem

Relative size, $\tau^a$	Marginal time (seconds) <sup>b</sup>	Accumulative time (seconds) <sup>c</sup>	Heuristic gap (%)
1	34.61	34.61	19.911
2	46.93	81.54	8.985
3	131.25	212.79	4.210
4	8,591.14	8,803.93	2.250
5	28,803.17	37,607.10	1.251
6	1,396.80	39,003.90	0.338
7	1,856.73	40,860.63	0.259
8	6,068.74	46,929.37	0.067
9	1,087.30	48,016.67	0.063
10	1,859.24	49,875.91	0.062
20	10,437.88	60,313.79	0.014
50	9,792.06	70,105.85	0.002
100	18,335.57	88,441.42	0.000



# Model Results (Case study: Annual planning for Qingdao)

- Cost savings and improved delivery distance
- Benchmark plan: it selects the best DS locations and block assignments from the first month without considering existing DSs and thus provides a theoretical lower bound on the total cost.

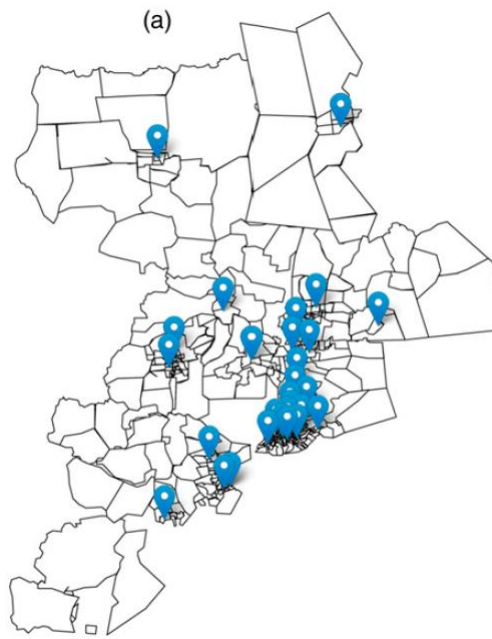
	Manual plan	Model plan (% change)	Benchmark plan (% change)
Total cost (CNY)	44,567,733	40,967,948 (−8.08%)	37,235,276 (−16.45%)
Fixed cost (CNY)	14,270,108	10,927,427 (−23.42%)	7,333,800 (−48.61%)
Delivery cost (CNY)	30,297,625	30,040,521 (−0.85%)	29,901,476 (−1.31%)
DS number in the last month	54	29 (−46.30%)	28 (−48.15%)
DS capacity use rate	37.14%	48.70% (11.56 pp)	74.18% (37.04 pp)
Average delivery distance (km)	5.53	5.33 (−3.61%)	5.22 (−5.58%)

8.08% total cost savings compared to manual plans

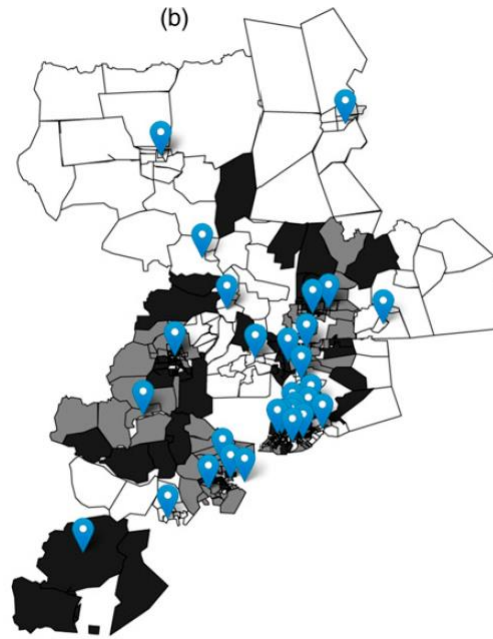
Reduced number of DSs, increased DS use rate

Shorter average delivery distance, improved customer experience

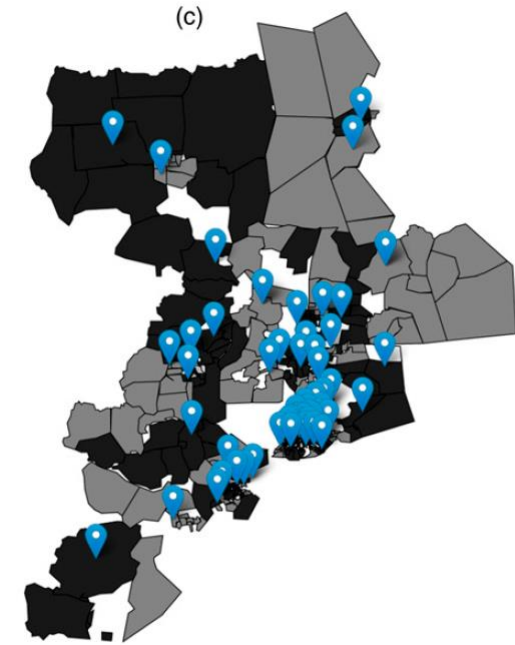
# Model Results (Case study: Annual planning for Qingdao)



Benchmark plan.



Model plan.

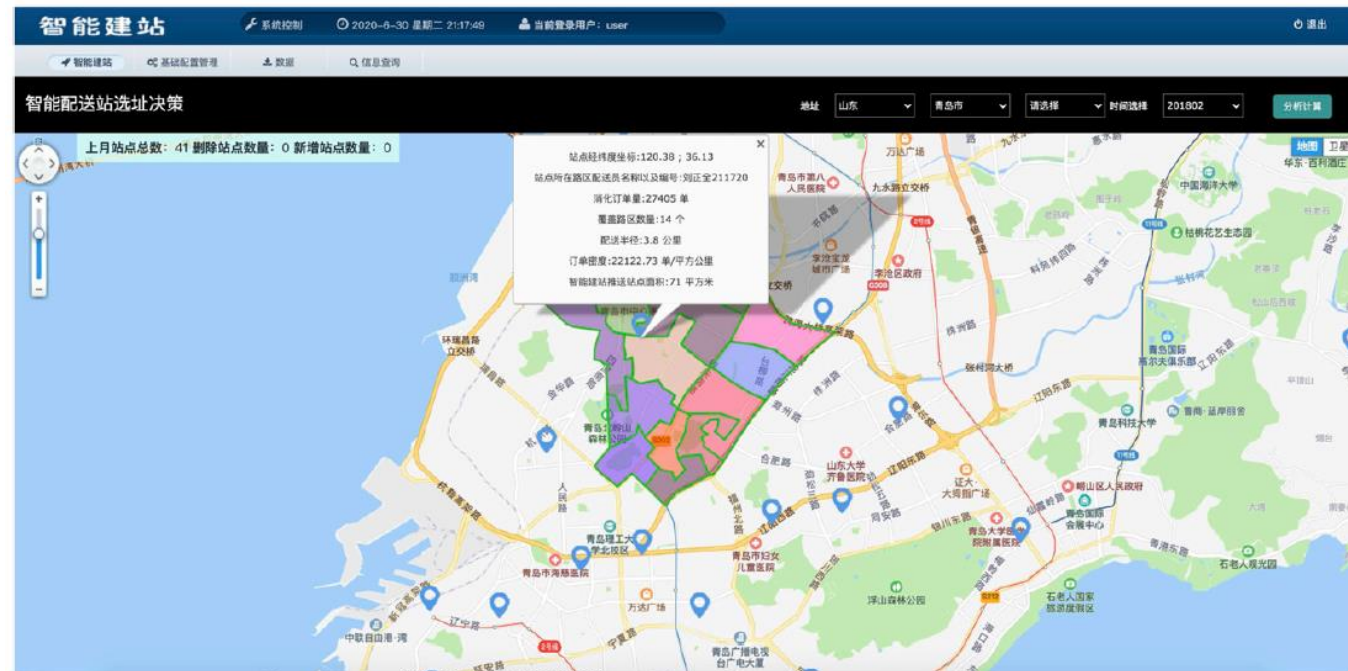


Manual plan.

Note: the grey and black blocks indicate differences compared to the benchmark plan.

# Implementation

- Embedded model in DS location planning tool
  - Tool features: historical data extraction, optimization, visual interface
  - Used in 16 major pilot cities





# Impact and Future Plans

- Estimated savings of over 82 million yuan in 2018
- Improved delivery timeliness and network efficiency
- Increased employee efficiency and faster planning process
  
- Challenges: housing resources, demand surges, updating plans
- Future plans: integrate real-time information, traffic data

# Quiz

- In a small town, there are three blocks (indexed by  $j$ ). The package delivery demand in the next four months (indexed by  $t$ ) is listed in the table below.

$j = \backslash t =$	1	2	3	4
1	30	40	50	10
2	20	50	20	60
3	50	40	50	20

- The delivery firm is planning the location of the single delivery station (DS). The cost of transportation is \$1 per package within a block and \$2 between blocks. Currently, the DS is in block 1. The cost of relocation is \$40. The operating cost of DS is zero. The minimum total cost of the next four months is \$740. True or false?