



Transit Network Design under Stochastic Demand

Hong K. LO

Civil and Environmental Engineering
Hong Kong University of Science and Technology



Background

➤ Public transport

- Fixed route transit (FRT) : fixed route and fixed schedule, large capacity, exclusive right of way, such as metro, bus, or regular ferry
- Demand responsive transit (DRT): flexible route, small or medium vehicle, complementary to FRT, such as ridesharing services, taxi, etc.
- Transit Network Design Problem (TNDP)
 - Find a most cost efficient operating plan for FRT while satisfying passenger demand
- Challenges in TNDP
 - Jointly optimize the FRT and DRT as an integrated system under stochastic demand, hard capacity constraints, and user equilibrium flows



Background

- Schedule-based TNDP
 - Ferry service network (Lai and Lo, 2004, Wan and Lo, 2009)
 - Time-space network
 - Decision variables: route, schedule, fleet size
 - Integer (a lot) and real variables
- Frequency-based TNDP
 - Rapid transit network (Bruno et al. 1998, Laporte et al. 2005, Samanta et al. 2011, Marin, 2007, Wan and Lo 2003, 2009)
 - Objective: minimize construction and passenger cost
 - Decision variables: route alignment, service frequency
 - Integer and real variables



Background

- Deterministic TNDP
 - Given OD matrix
 - First formulated by Magnanti (1984) as a mixed integer linear program (MILP)
- Stochastic TNDP
 - Stochastic programming approach (Ruszczynski, 2008)
 - Stochastic demand follows a known probability distribution
 - Monte-Carlo simulation to approximate the cost expectation by sample average, two-stage stochastic problem
- Robust optimization approach
 - Stochastic demand captured by an uncertainty set
 - Min-max problem, worst case scenario (Ben-tal et al., 2004)₄



Background

- Solution algorithm for stochastic approach
 - Exact method: multi-dimensional integral evaluation, formidable task
 - Heuristic approach: search the neighborhood of the initial solution, efficient but solution quality not guaranteed (Hoff et al. 2011)
 - Approximation method: L-shaped/ Multi-cut method, long computation time, global optimal solution is not guaranteed
- Solution algorithm for robust optimization approach
 - Polyhedral uncertainty set, linearization
 - Same dimension as its deterministic counterpart
 - Conservative, dependent on the size of the uncertainty set



Objectives

- To develop a modeling framework for combining FRT and DRT network design under stochastic demand
 - Investigate the benefits of the integrated services under stochastic demand
 - Develop a **service reliability (SR)** based formulation and solution algorithm to address demand uncertainty
- To assess the performance of SR-based formulation
 - Two application contexts: ferry network and rapid transit network
 - Two passenger flow distribution pattern: system optimal (SO) and user equilibrium (UE)



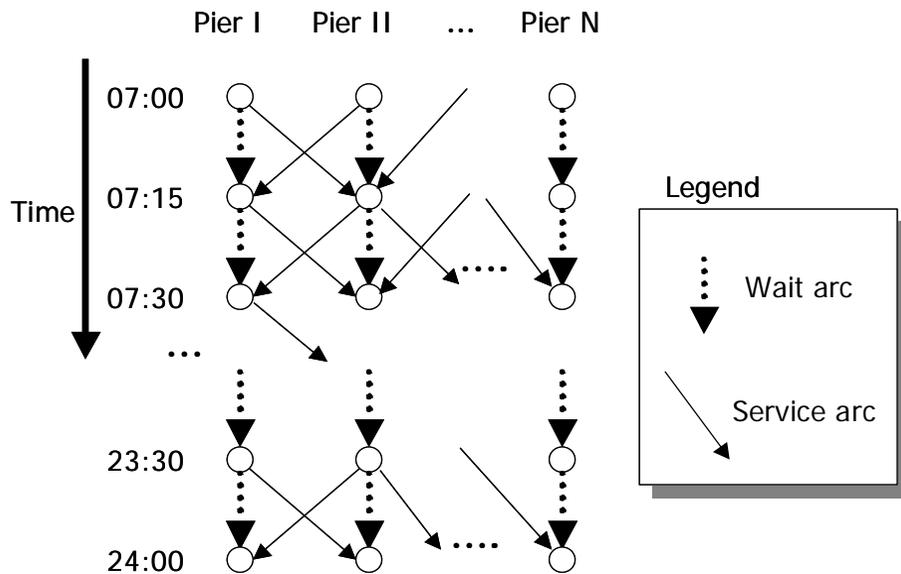
Transit Network Design with Stochastic Demand under System Optimal Flows

Lo, H., K. An and W. Lin. 2013. Ferry Network Design under Demand Uncertainty. **Transportation Research Part E**, 59, 48-70.

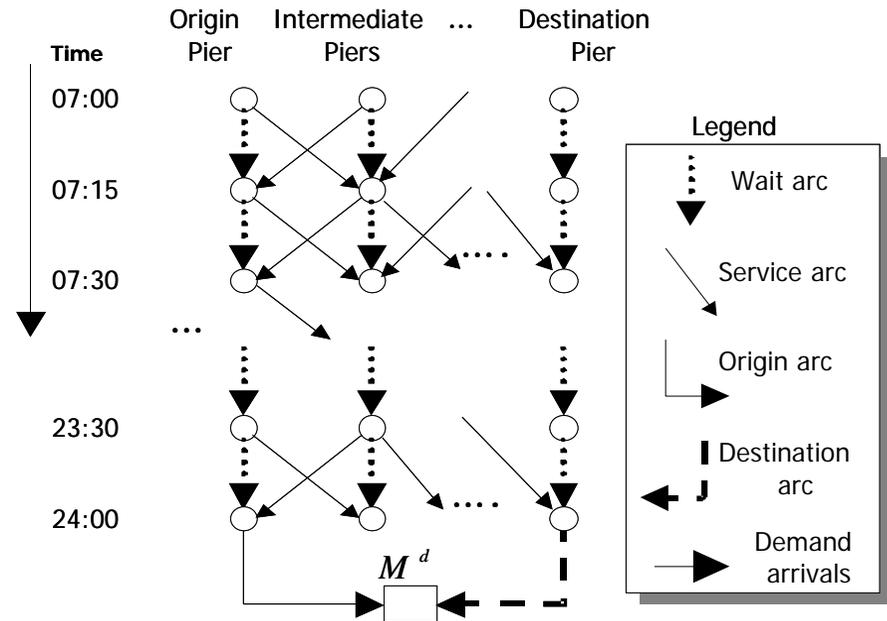


Time-space network description

➤ Ferry time-space network



➤ Passenger time-space network



➤ Notations

$$\mathbf{Y} = \{Y_{ij}\}, ij \in \begin{cases} S^f \\ W^f \end{cases}$$

Ferry service arc
Ferry waiting arc

$$\mathbf{X} = \{X_{ij}^d\}, ij \in \begin{cases} S^d \\ W^d \end{cases}$$

Passenger service arc
Passenger waiting arc



Research objective

Regular services



Ad-hoc services



Regular ferry services (FRT)

Fixed schedule

Large capacity

Low unit cost

+

Ad-hoc ferry services (DRT)

Flexible schedule

Small capacity

High unit cost

➤ Objectives

➤ Capture demand uncertainty on ferry service deployment

➤ Develop a modeling framework for combining FRT and DRT

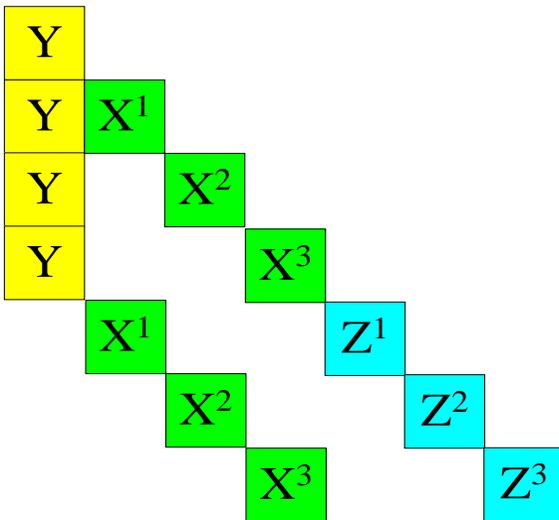


Problem challenge

➤ Objective:



➤ Decisions:



The original problem

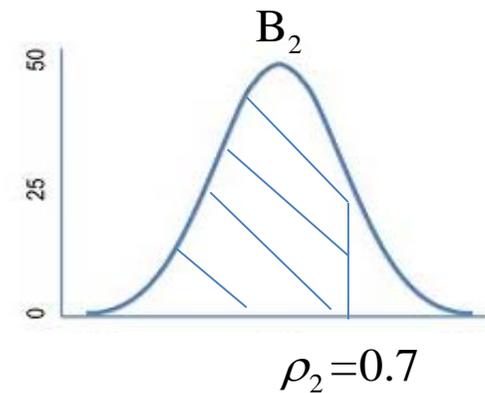
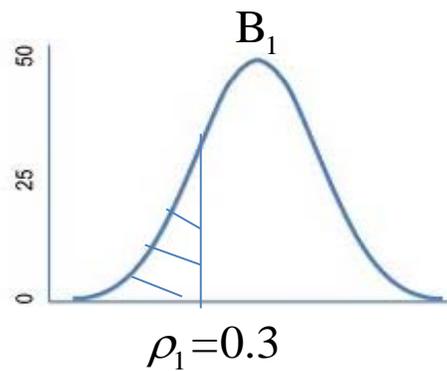
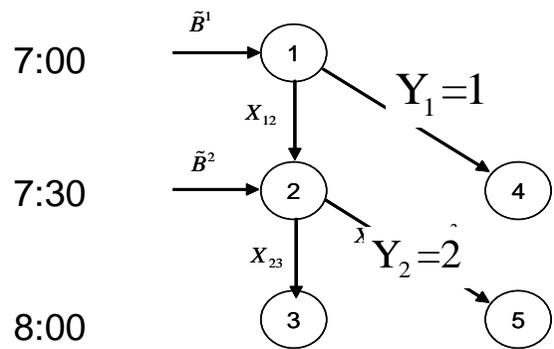
- The deployment of these two services are related
- Two-stage stochastic program
- Represent stochastic demand by a large number of discrete scenarios
- A large size MILP



Service reliability ρ

- The probability of passengers carried by regular ferry services
- A vector, one for an OD pair

	Service reliability	Regular service	Ad-hoc service
Tradeoff	high	more	less
	low	Less	more





SR-based stochastic formulation (Phase-1)

Phase-1 regular service deployment **Y**

$$\min_{\mathbf{Y}} \sum_{i \in N_b^f} \sum_{j \in N^f \setminus N_b^f} Y_{ij} c_F^1 + \sum_{ij \in S^f} Y_{ij} c_{ij}^1$$

Fixed cost

Operating cost

1

(1)
$$\sum_{j \in N^d} X_{ij}^d - \sum_{k \in N^d} X_{ki}^d = \begin{cases} \bar{B}^d & \text{if } i \text{ is the O of } OD \ d \\ -\bar{B}^d & \text{if } i \text{ is the D of } OD \ d \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in N^d$$

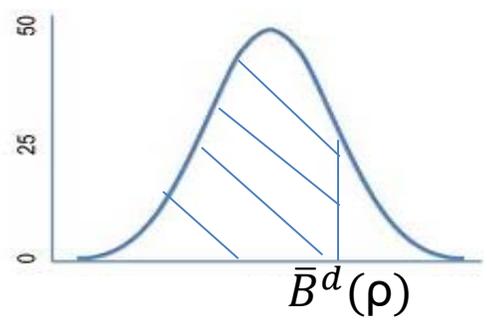
Passenger flow conservation

(2)
$$\bar{B}^d = \Psi_d^{-1}(\rho^d)$$

$$0 \leq \rho^d \leq 1 \quad \forall d$$

Service reliability constraint

Stochastic demand



2

Demand realized within the service reliability boundary can be carried by regular ferry services



SR-based stochastic formulation (Phase-1)

$$\mathbf{Y} \quad \min_{\mathbf{Y}} \quad \sum_{i \in N_b^f} \sum_{j \in N^f \setminus N_b^f} Y_{ij} c_F^1 + \sum_{ij \in S^f} Y_{ij} c_{ij}^1$$

Decision variable

Fixed cost

Operating cost

(3)

$$\sum_{j \in N} Y_{ij} - \sum_{k \in N} Y_{ki} = 0 \quad \forall i \in N$$

$$0 \leq Y_{ij} \leq U_{ij}$$

$$Y_{ij} \in \text{integer} \quad \forall ij \in A$$

Regular services connection constraints

(4)

$$\sum_{i \in N_b^f} \sum_{j \in N^f \setminus N_b^f} Y_{ij} \leq V$$

Fleet size constraint

(5)

$$\sum_{d \in R} X_{ij}^d \leq Y_{ij} \zeta \quad \forall ij \in S^f$$

Capacity constraints



SR-based stochastic formulation (Phase-2)

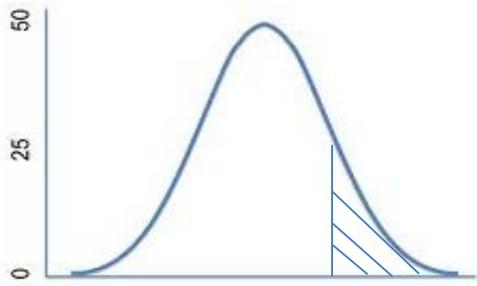
Phase-2 Ad-hoc services deployment Z_e

$$Z_e \min_{Z_e} \bar{Q}(\rho) = \sum_{e \in E} p_e \left(\underbrace{\sum_{d \in R} c_d^3 Z_e^d}_{\text{Ad-hoc cost}} + \underbrace{\sum_{d \in R} \sum_{ij \in A} c_{ij}^2 X_{ij,e}^d}_{\text{Passenger cost}} \right)$$

1

- (1)
$$\sum_{j \in N} X_{kj,e}^d - \sum_{i \in N} X_{ik,e}^d \begin{cases} = B_e^d - Z_e^d, & \text{if } k \text{ is the O of OD } d \\ = Z_e^d - B_e^d, & \text{if } k \text{ is the D of OD } d, \forall k \in N, d \in R \\ = 0 & \text{otherwise} \end{cases}$$
 Passenger flow conservation
- (2)
$$\sum_{d \in R} X_{ij,e}^d \leq Y_{ij} \zeta \quad \forall ij \in S^f$$
 Capacity constraints

Stochastic demand



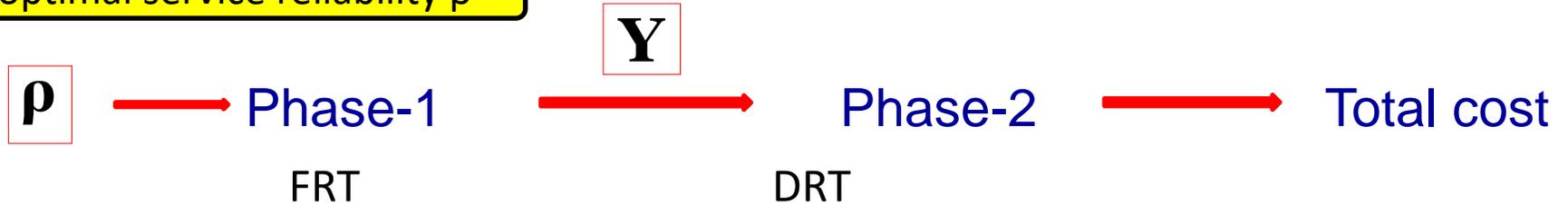
2

Demand beyond the service reliability is to be carried by ad-hoc services

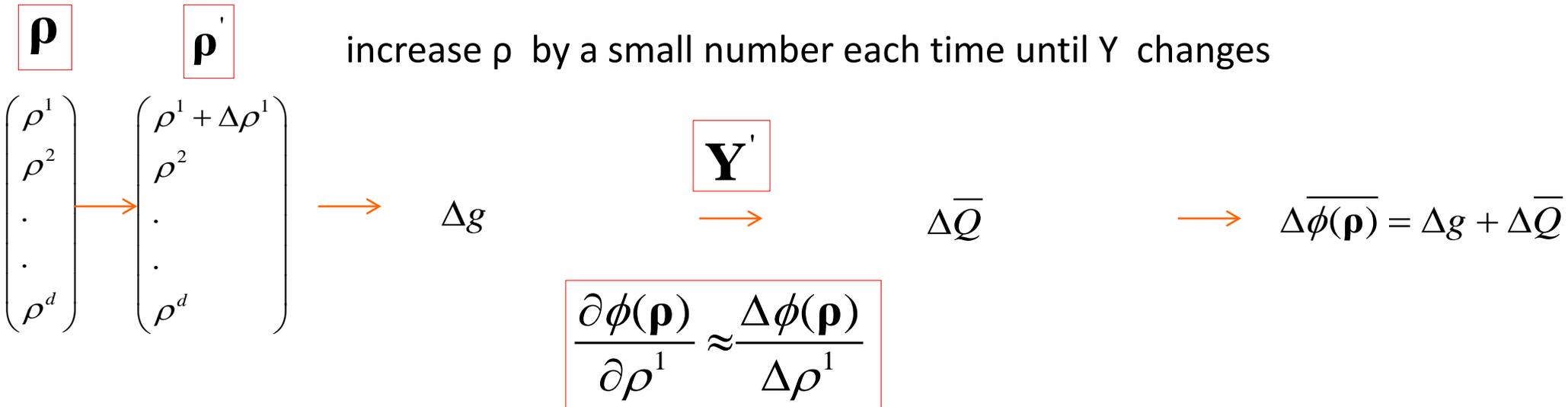


SR-based gradient approach

Find the optimal service reliability ρ



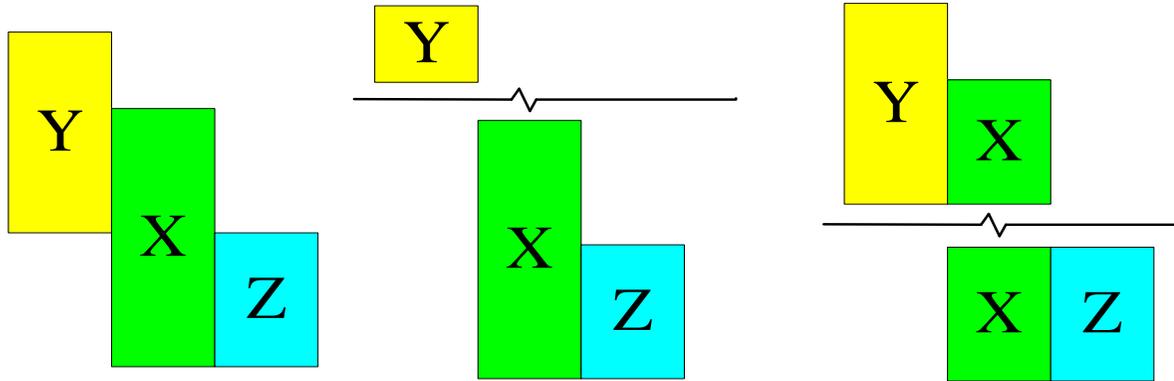
Find the gradient of total cost w.r.t ρ



Disadvantage: computation time highly depends on the number of OD pairs d



Advantage of the SR-based model



The original problem L-shaped method SR-based gradient method

The constraint structures

- Take advantage of the special structure of the problem
- Separate the large size MILP into one smaller size MILP and one LP
- Can be extended to include the user equilibrium assignment principle



Transit Network Design with Stochastic Demand under User Equilibrium Flows

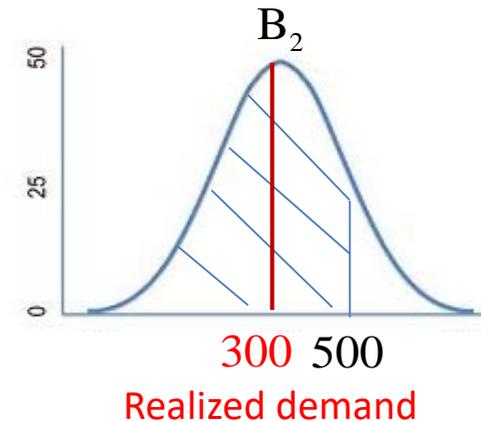
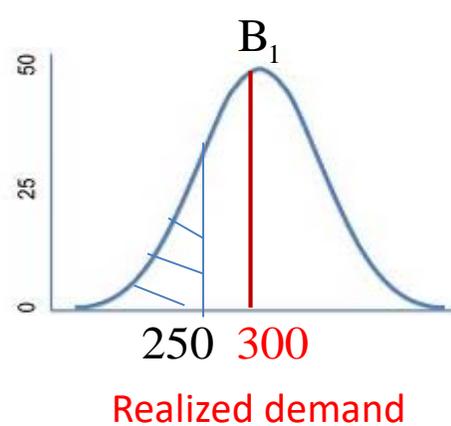
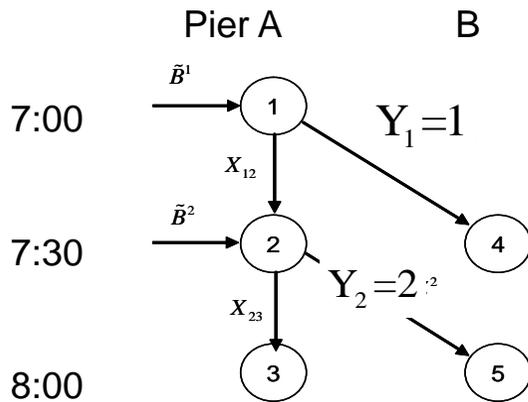
An, K. and H. Lo. 2014. Ferry Service Network Design with Stochastic Demand under User Equilibrium Flows. **Transportation Research Part B**, 66, 70-89.



Scheme A (with passenger reservation)

➤ Allow advance reservation for passengers

- Passenger demand for the next day is obtained one-day in advance
- Provide sufficient amount of ad-hoc services considering user equilibrium

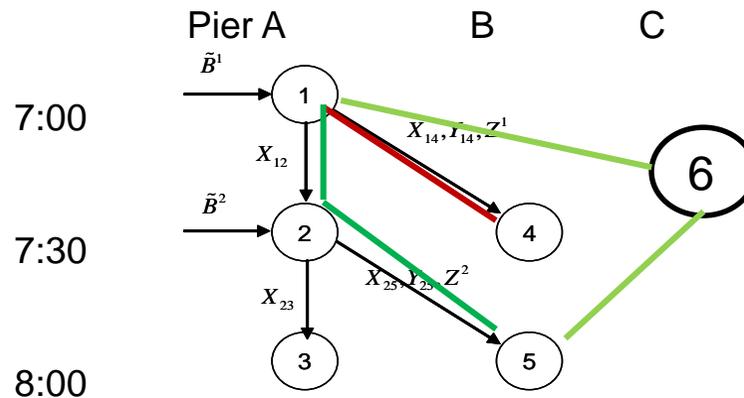


Demand realization $b_1=300$, $b_2=300$, regular service capacity= 250, 500
no need to provide ad-hoc services, reduce operating cost



Scheme A (with passenger reservation)

- User equilibrium with stochastic demand
 - Not sensible to find the long-term UE as demand varies from day to day
 - Find the short-term UE : for a demand realization, same minimum traveling cost for passengers on the same origin-destination (OD)
 - Passenger options: take the congested direct service, wait for the next direct service or take a detour



Travelling cost= waiting + on vehicle + overflow delay



Stochastic Formulation under UE (Phase-1)

Phase-1 Regular ferry services routes and schedule **Y**

The same as SO

$$\min_{\mathbf{Y}} \sum_{i \in N_b^f} \sum_{j \in N^f \setminus N_b^f} Y_{ij} c_{ij}^1 + \sum_{ij \in S^f} Y_{ij} c_{ij}^1$$

Regular services cost

input

$$\sum_{j \in N^d} X_{ij}^d - \sum_{k \in N^d} X_{ki}^d = \begin{cases} \bar{B}^d & \text{if } i \text{ is the O of } OD \ d \\ -\bar{B}^d & \text{if } i \text{ is the D of } OD \ d \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in N^d$$

- 1) Regular services connectivity constraints
- 2) Passenger flow conservation constraints
- 3) Capacity constraints



Stochastic Formulation under UE (Phase-2)

Phase-2 Dynamic ad-hoc services deployment Z_e e for day scenario

$$\min_{Z_e} \bar{Q}(\rho) = \sum_{e \in E} p_e \left(\underbrace{\sum_d c_d^3 Z_e^d}_{\text{Ad-hoc cost}} + \underbrace{\sum_{ij \in A} c_{ij}^2 \sum_d X_{ij,e}^d}_{\text{Travel time}} + \underbrace{\sum_{ij \in A} -\beta_{ij,e} \left(\sum_d X_{ij,e}^d \right)}_{\text{Overflow delay}} \right)$$

Passenger cost

(P4.2) For a certain ad-hoc services cost $\theta_{e,\min} \leq \theta_e \leq \theta_{e,\max}$

$$\min_{X,Z} f = \sum_{ij \in A} c_{ij}^2 \sum_d X_{ij,e}^d \quad \text{Travel time cost}$$

$$\sum_{j \in N} X_{kj,e}^d - \sum_{i \in N} X_{ik,e}^d = \begin{cases} B_e^d - Z_e^d, & \text{if } k \text{ is the origin of OD } d \\ 0 & \text{otherwise} \end{cases}, \forall k \in N$$

$$\sum_d X_{ij,e}^d \leq \zeta Y_{ij}, \quad \forall ij \in A \quad \xrightarrow{\beta_{ij,e}} \quad \sum_{ij \in A} -\beta_{ij,e} \left(\sum_d X_{ij,e}^d \right) \quad \text{Overflow delay}$$

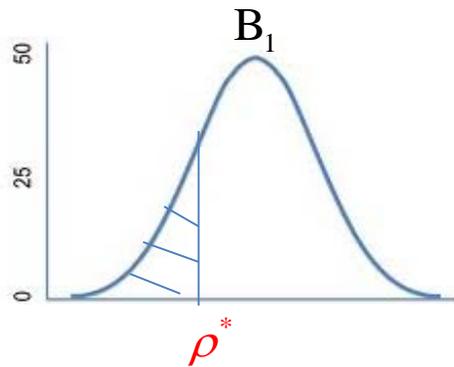
$$\sum_{d \in D} c_d^4 Z_e^d = \theta_e \quad \text{input} \quad \text{Ad-hoc cost}$$

Proposition 4.1: P4.2 yields a UE flow pattern under capacity constraints, with the negative Lagrange multiplier associated with the link capacity constraint representing the corresponding passenger overflow delay.

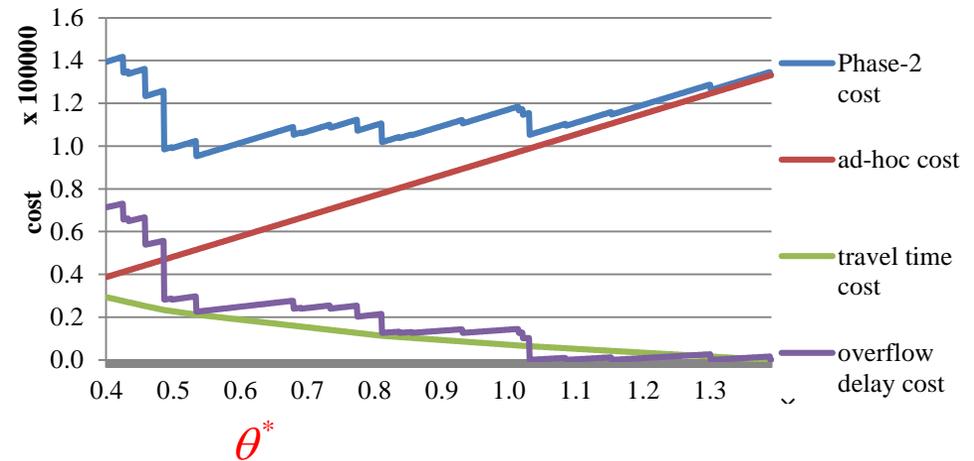


SR-based gradient solution procedure

- **Phase-1:** Regular ferry services routes and schedule
- **Phase-2:** Dynamic ad-hoc deployment
- Total cost= Phase-1 + Phase-2 cost



Find the optimal service reliability



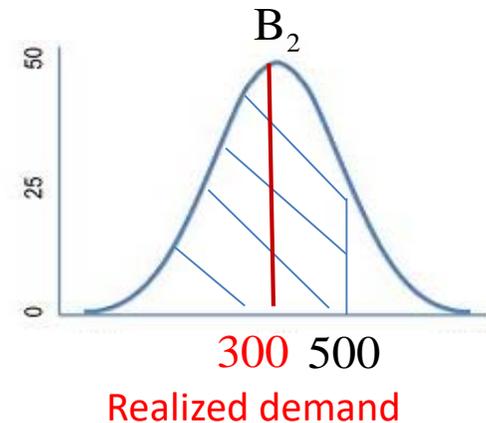
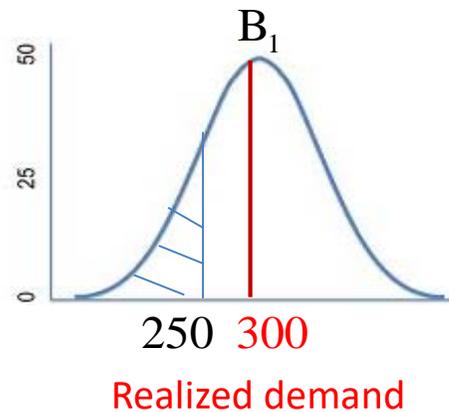
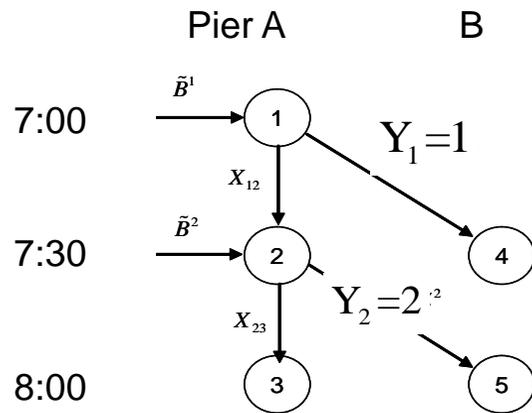
Find the optimal ad-hoc services provision level

- Increase θ by κ each time and the one with the lowest cost is kept as the optimal solution



Scheme B (no advance reservation)

- Passenger demand is revealed only when they arrive at the piers
- Provide ad-hoc services whenever there is demand overflow

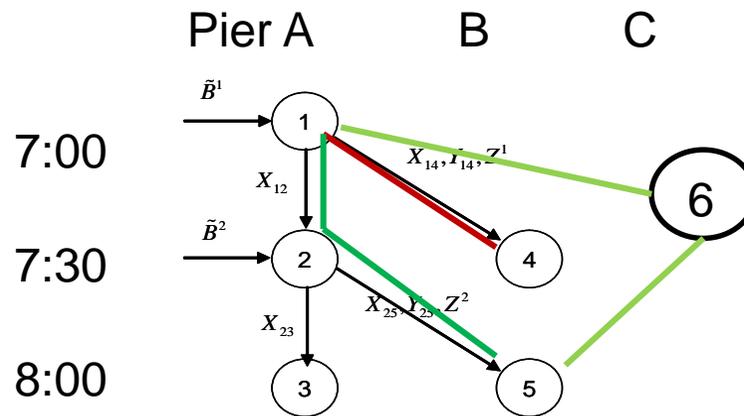


Demand realization $b_1=300$, $b_2=300$, provide ad-hoc services 50, spaces wasted for B_2



Scheme B (no advance reservation)

- No waiting or detour, only direct service, best for passengers
- Most costly plan for the company



Travelling time = on vehicle travel time



SR-based stochastic formulation (single phase)

- **Phase-1:** Regular ferry services routes and schedule \mathbf{Y}
- Up to demand $\underline{B}_d = \Psi_d^{-1}(\rho^d)$
- **Phase-2:** Dynamic ad-hoc deployment $Z_{d,e} = \max\{0, B_{d,e} - \underline{B}_d\}$
- Expected ad-hoc services cost can be calculated by $\bar{\theta}(\boldsymbol{\rho}) = \sum_d \int_{B_{d,e} \geq \underline{B}_d} c^3 Z_{d,e} \Psi_d de$
- **Single phase problem**

$$\min_{\mathbf{h}, \mathbf{Y}, \boldsymbol{\rho}} \phi(\boldsymbol{\rho}) = (\mathbf{c}^1)^T \mathbf{Y} + \bar{\theta} + (\mathbf{c}^4)^T \mathbf{t}^T \bar{\mathbf{B}}$$

Regular service cost	Ad-hoc service cost	Passenger cost
----------------------------	---------------------------	----------------

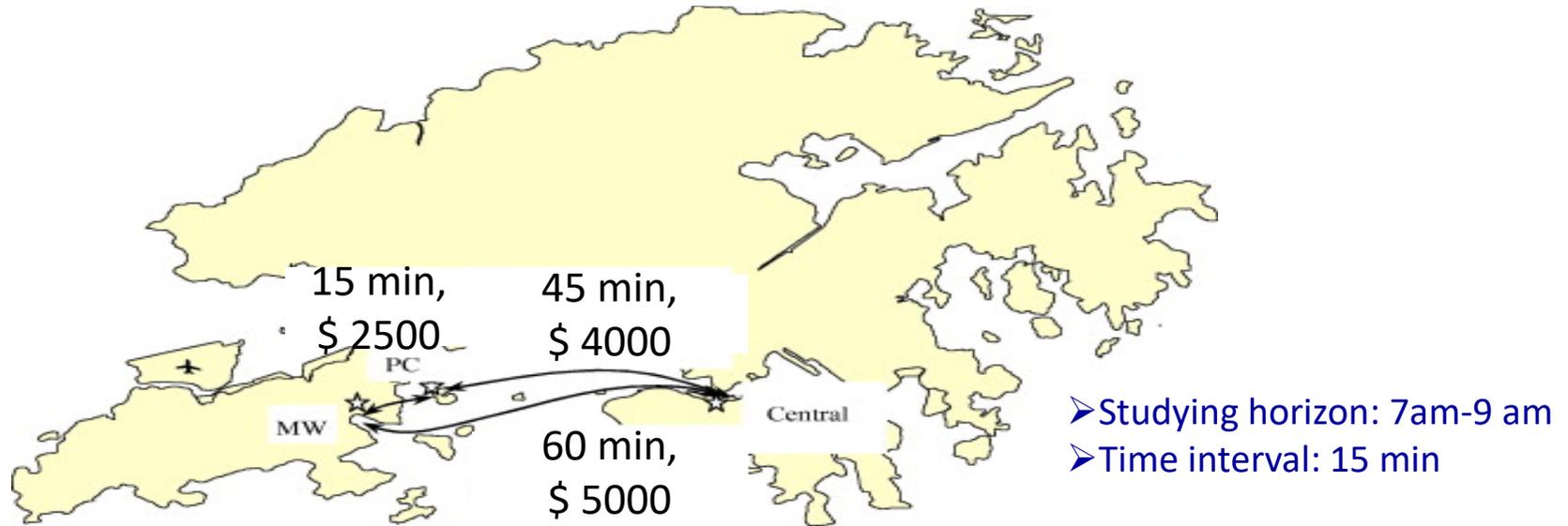
s.t. Regular services connection constraints
 Fleet size constraint
 Capacity constraints
 Passenger flow conservation

- **Value of reservation:** cost difference between scheme A and B



Numerical study

➤ Ferry service network in Hong Kong

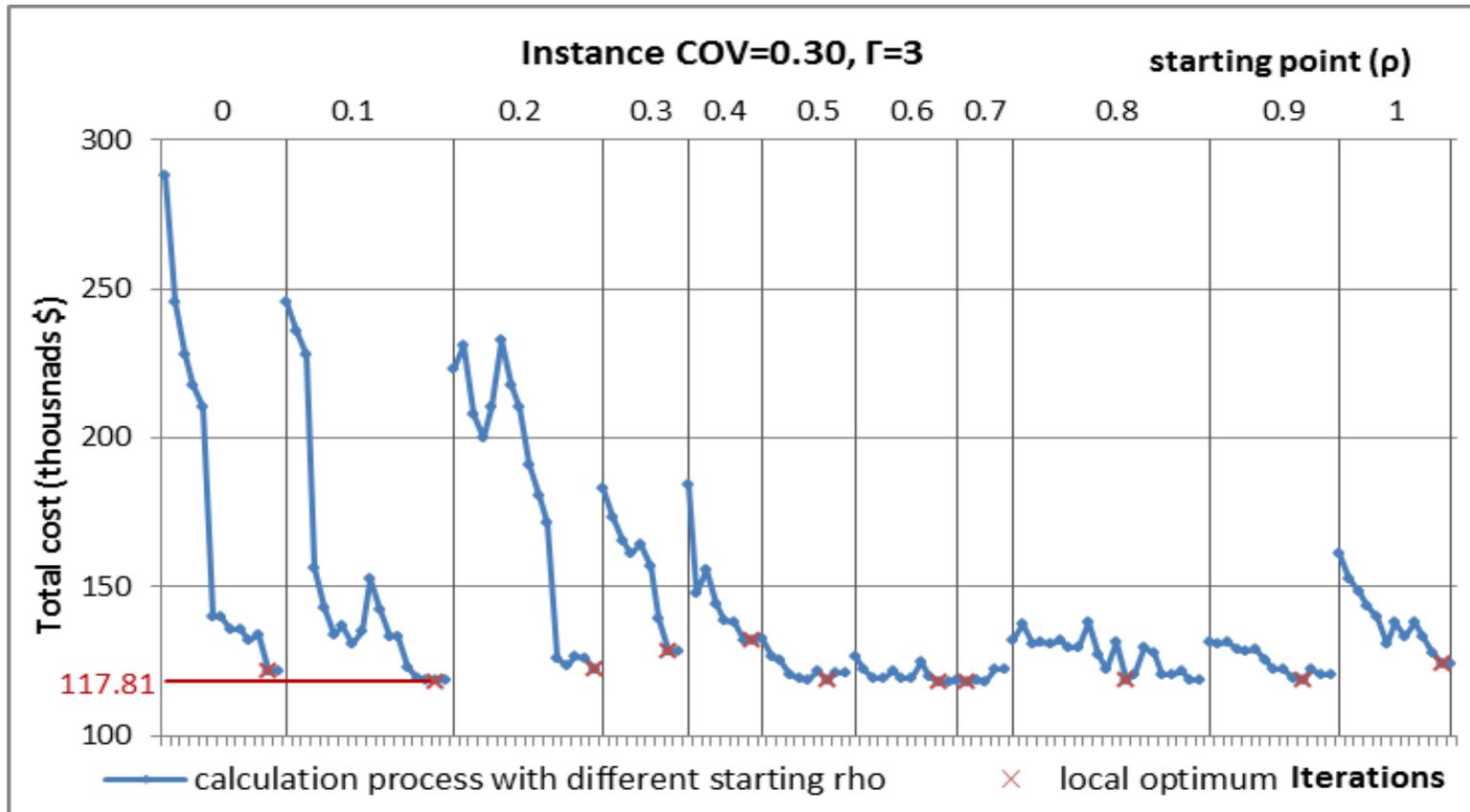


Time Slice	CBD-MW	MW-CBD	CBD-PC	PC-CBD	MW-PC	PC-MW
1	80	0	60	0	0	0
2	250	45	230	20	30	75
3	420	100	400	80	80	100
4	350	140	330	140	140	180
5	200	130	180	130	190	230
6	100	120	80	120	320	170
7	90	110	70	70	210	70
8	50	50	30	50	100	50



Solution procedure illustration

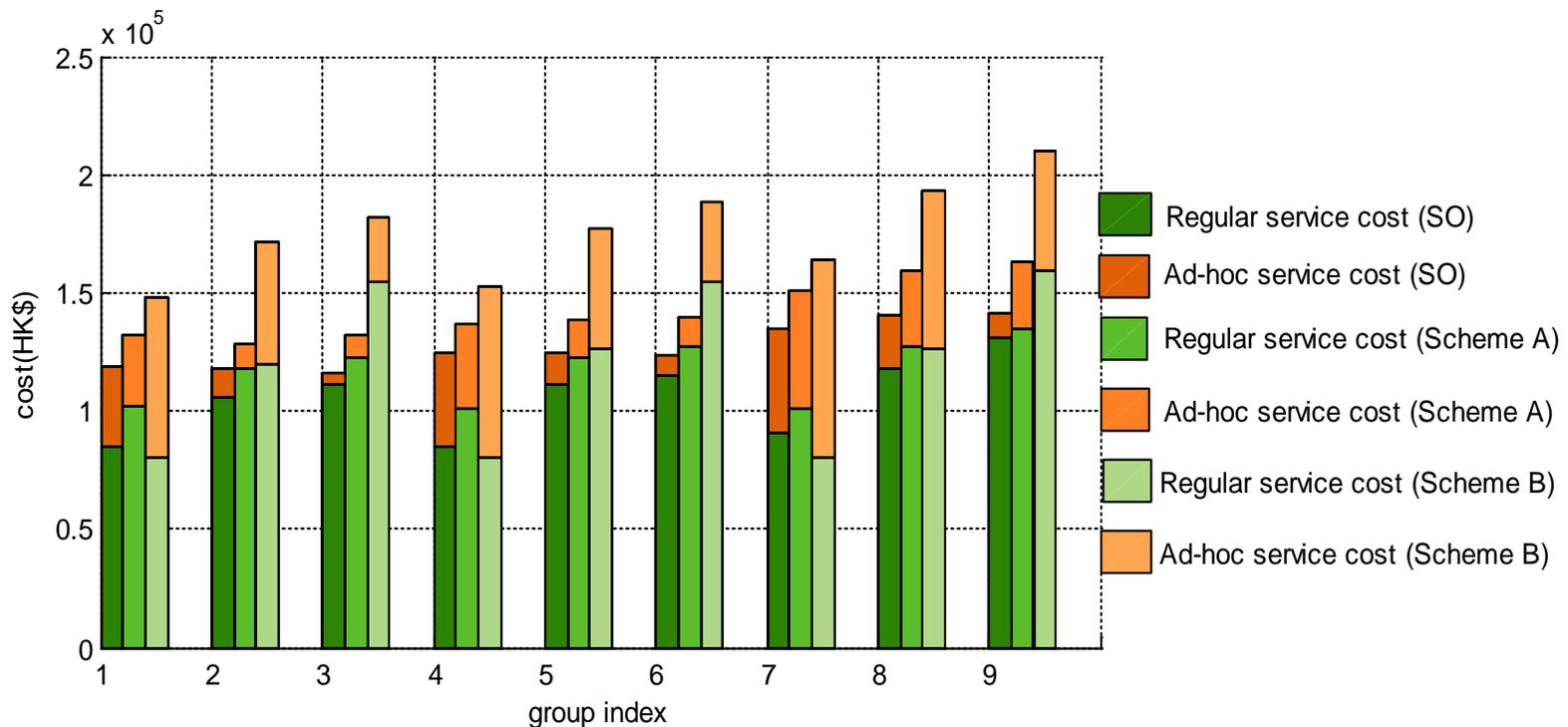
- Different starting points comparison
- Computation time: SR-based method : 217 seconds versus L-shaped method: 6 hours





Result analysis

- Group: different unit ad-hoc cost and COV (coefficient of variation)
- Total company cost under SO is lower than that under UE
- The value of reservation between scheme A and B increase with COV



Service deployment comparison between UE and SO solutions



Robust Rapid Transit Network Design

An, K. and H. Lo. 2016. Two-Phase Stochastic Program for Transit Network Design under Demand Uncertainty. **Transportation Research Part B**, 84, 157-181.

An, K. and H. Lo. 2015. Robust Transit Network Design with Stochastic Demand Considering Development Density. 21th International Symposium on Traffic and Transportation Theory (ISTTT) and **Transportation Research Part B**, 81, 737-754.



Research objectives

- Rapid TNDP: location of stations, route alignment and frequency



+

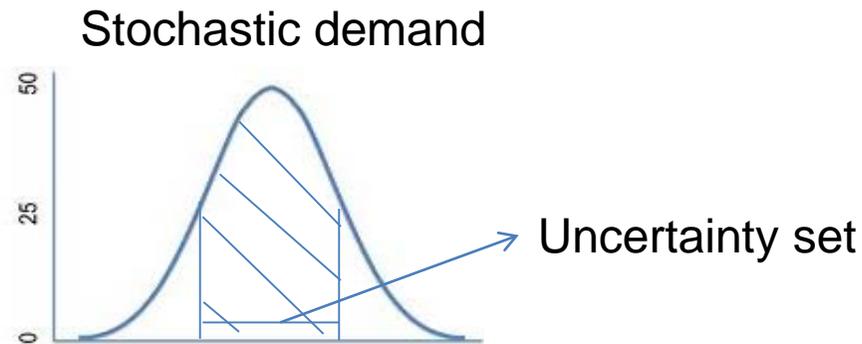


- Objectives
 - Investigate the TNDP by robust optimization
 - The optimal level of robustness to minimize the system cost
- Methodology
 - Apply the SR-based two-phase formulation
 - Improve the efficiency of the gradient solution procedure



Robust optimization to address demand uncertainty

- Robust optimization
 - Day to day variation in demand, min-max problem, worst case scenario (Ben-tal et al., 2004)
 - Conservative solutions
- Evaluate the outcome of uncertainty set
 - Search for the optimal robustness level to hedge against uncertainty
 - Demand fluctuation, uneconomical to rely on transit lines alone

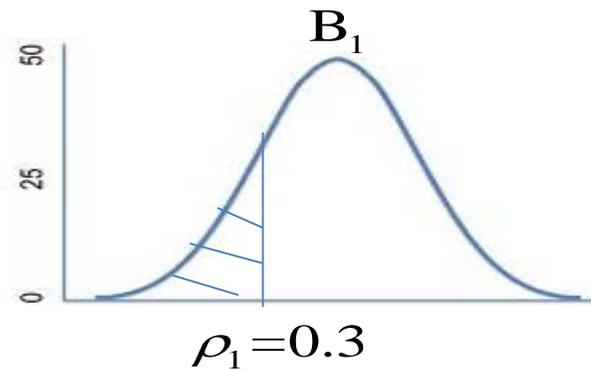


system cost vs. robustness level



Service reliability to address demand uncertainty

- Service reliability ρ : the probability that passengers can be carried by rapid transit services
- Conveys the level of robustness, size of the uncertainty set
- Dial-a-ride services cost evaluate the outcomes beyond the uncertainty set



Rapid transit services

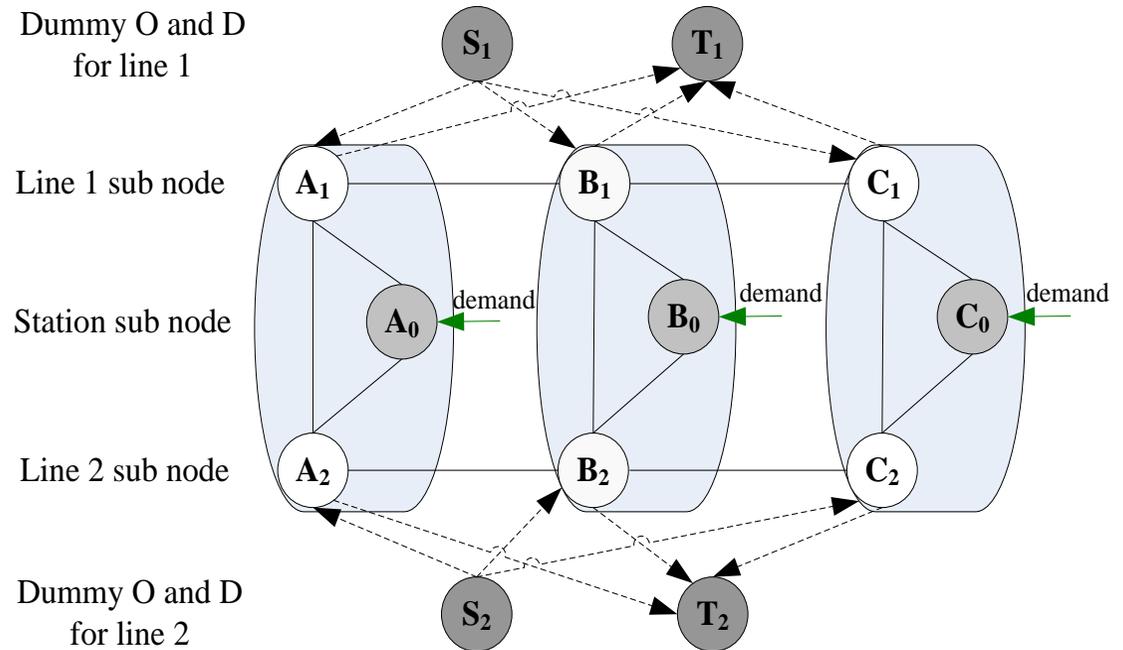
+

Dial-a-ride services



Network representation

- Multi-line design: dummy origin and destination
- Passenger transfer and waiting: line sub node and station sub node
- Dummy O and D: dark gray node
- Dummy arcs: black dashed line
- Transfer time: blue line
- Get on/off time: green line
- On vehicle travel time: orange line
- Demand: imposed on the station sub node





Two-phase Robust Formulation under System Optimal Flows



Two phase robust formulation

Phase-1

Rapid Transit line (RTL) alignment and frequencies $\mathbf{W}, \mathbf{Y}, \mathbf{f}$

$$\min_{\mathbf{W}, \mathbf{f}, \mathbf{Y}} \sum_{r \in R} \sum_{ij \in A} c_{ij}^1 f_r Y_{ij}^r + \sum_{r \in R} \sum_{ij \in A} c_{ij}^2 Y_{ij}^r + \sum_{i \in N} c^3 W_i$$

Transit operating

Line
construction

Station
construction

Phase-2

Dynamic dial-a-ride services deployment \mathbf{Z}_e

$$\min_{\mathbf{Z}_e} \bar{Q}(\boldsymbol{\rho}) = \sum_{h \in H} p_e \left(\sum_{ij \in A} c_d^3 Z_e^d + \sum_{ij \in A} c_{ij}^4 \sum_{d \in D} X_{ij,e}^d \right)$$

Dial-a-ride cost

Passenger cost

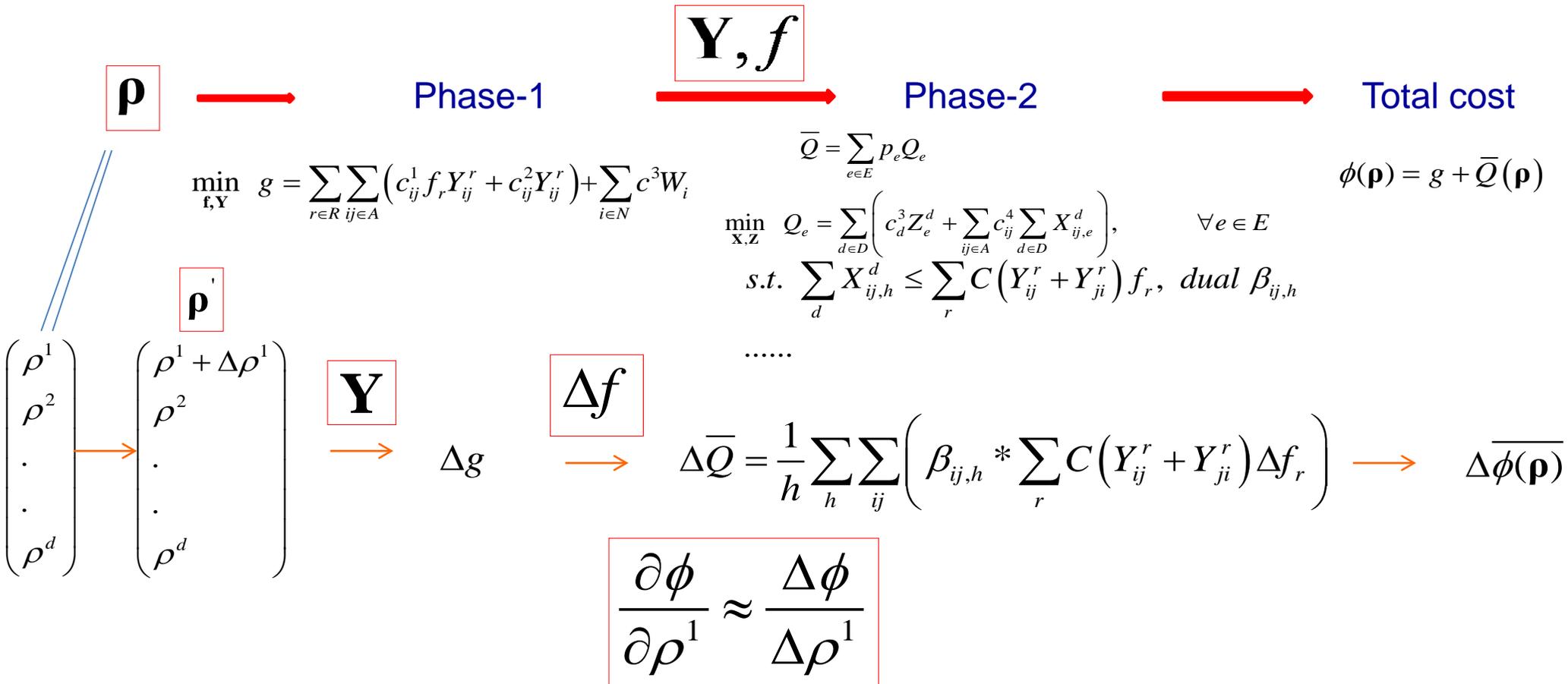
Decision variables:

- \mathbf{W} : Station location
- \mathbf{Y} : Route alignment
- \mathbf{f} : Frequency
- \mathbf{Z} : Dial-a-ride services

➤ Total cost = Phase-1 + Phase-2 cost



Solution algorithm : find the decent direction (revised)



Advantages: computation time does not depend on the number of OD pairs
 Assumption: A small perturbation in ρ changes frequency f only while maintaining the line alignment Y



Two Phase Robust Formulation under User Equilibrium Flows



Two-phase robust formulation

- Short term UE with stochastic demand
 - Passenger options: different routes choices on transit line
 - Over flow delay: transit services operated at capacity
 - Passenger travelling cost= on-vehicle travel time+ overflow delay

Phase-1: Feasible region for $\rho \in [0, 1]$

Phase-2: Dial-a-ride provision $\theta_e \in [\theta_{e,\min}, \theta_{e,\max}]$, e for one demand realization

External variables

(P6.4): Total required dial-a-ride services $\theta_{e,\min}$ just enough to carry demand overflow

Lower bound
$$\min_{\mathbf{Z}} \theta_{e,\min} = (\mathbf{c}^3)^T \mathbf{Z}_e$$

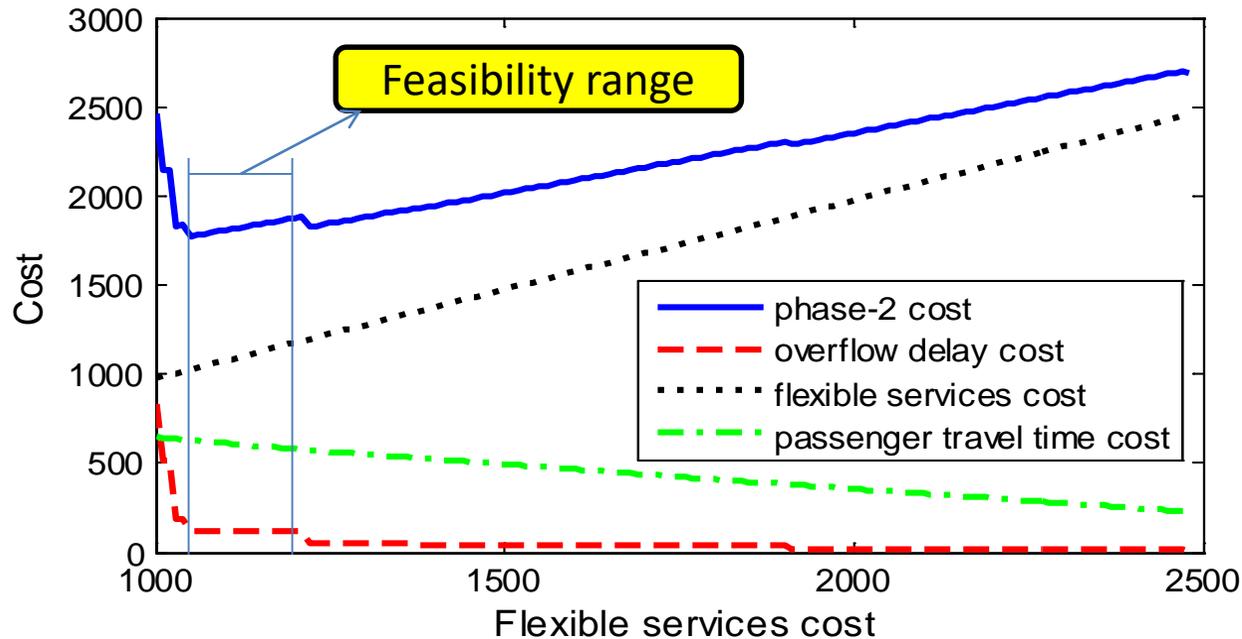
(P6.2) Maximum dial-a-ride services cost $\theta_{e,\max}$

Upper bound When the overflow delay is zero, every passenger can take the shortest path



Solution procedure

(P6.3): Find the optimal dial-a-ride services provision level $\theta_{e,\min} \leq \theta_e \leq \theta_{e,\max}$



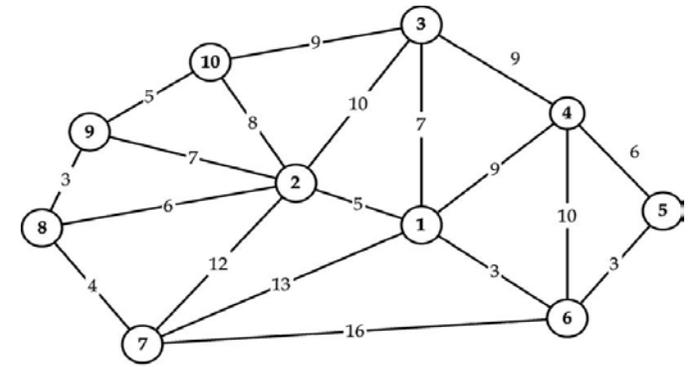
Detect the feasibility range: $[\underline{\theta}_e, \overline{\theta}_e]$

Advantage: (1) be able to find the exact optimal solution

(2) reduce the computation time to 1/10

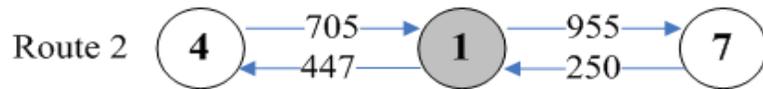


Case study

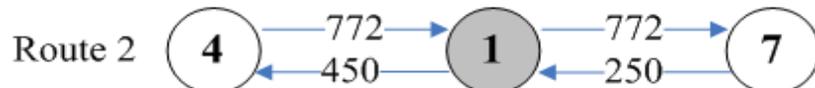
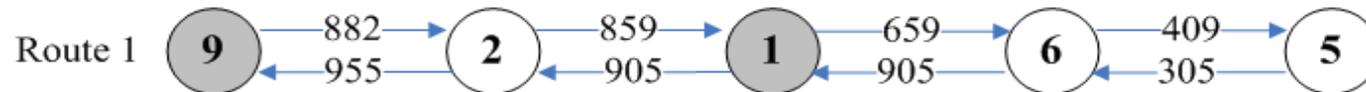


➤ Three lines were generated

➤ Frequency SO: $f_1 = 16, f_2 = 13, f_3 = 7$ vehicle / h



➤ Frequency UE: $f_1 = 19, f_2 = 16, f_3 = 7$ vehicle / h





Result analysis

- RTL carries most of the passengers while dial-a-ride serves as a supplementary role
- UE requires more dial-a-ride services to bring down the overflow delay cost

		SO				UE				
OD pair	OD demand	Path by RTL	Expected RTL Patronage	Transfer	Distance by RTL	Path by RTL	Expected RTL Patronage	Transfer	Distance by RTL	Distance
2→10	200	2-10	200	0	8	2-9-10	50	1	12	8
3→2	150	3-2	150	0	10	3-10-9-2	23	1	21	10
4→7	800	4-1-7	705	0	22	4-1-7	772	0	22	22
5→8	350	5-6-1-2-8	350	0	17	5-6-1-2-9-8	305	1	21	17
6→9	600	6-1-2-8-9	524	0	17	6-1-2-9	600	0	15	15
7→6	250	7-1-6	250	1	16	7-1-6	250	1	16	16
8→3	400	8-2-3	355	1	16	8-9-10-3	368	0	17	16
9→4	450	9-8-2-1-4	447	1	23	9-2-1-4	450	1	21	21
10→5	500	10-2-1-6-5	393	1	19	10-9-2-1-6-5	409	1	23	19

91 %

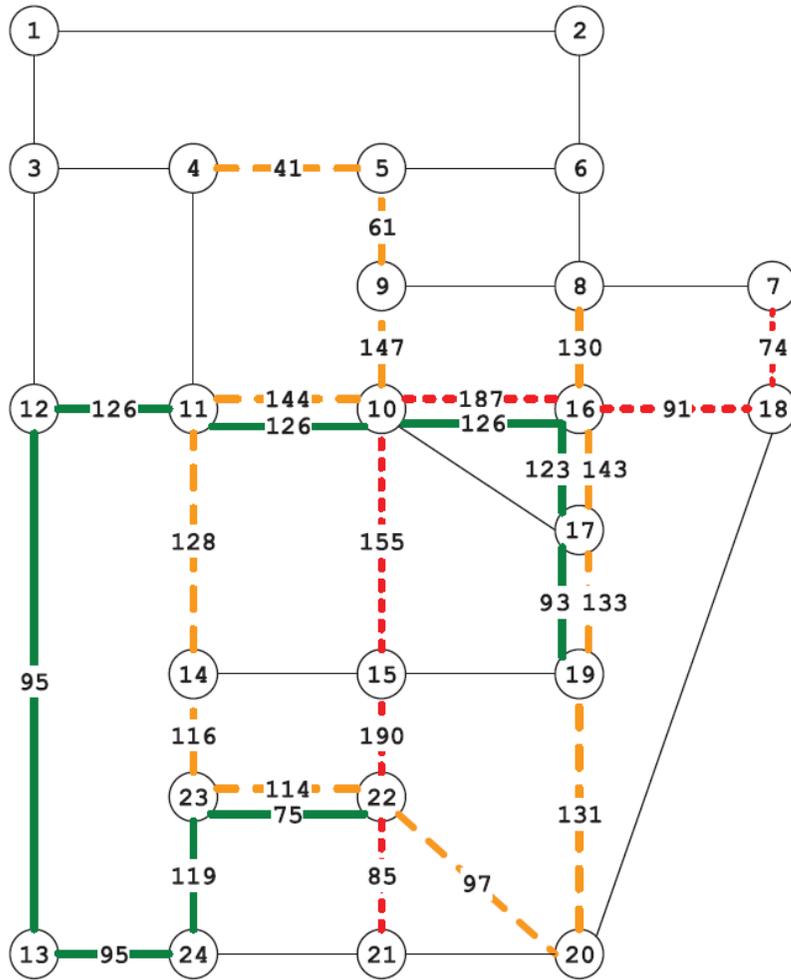
87 %



Result analysis

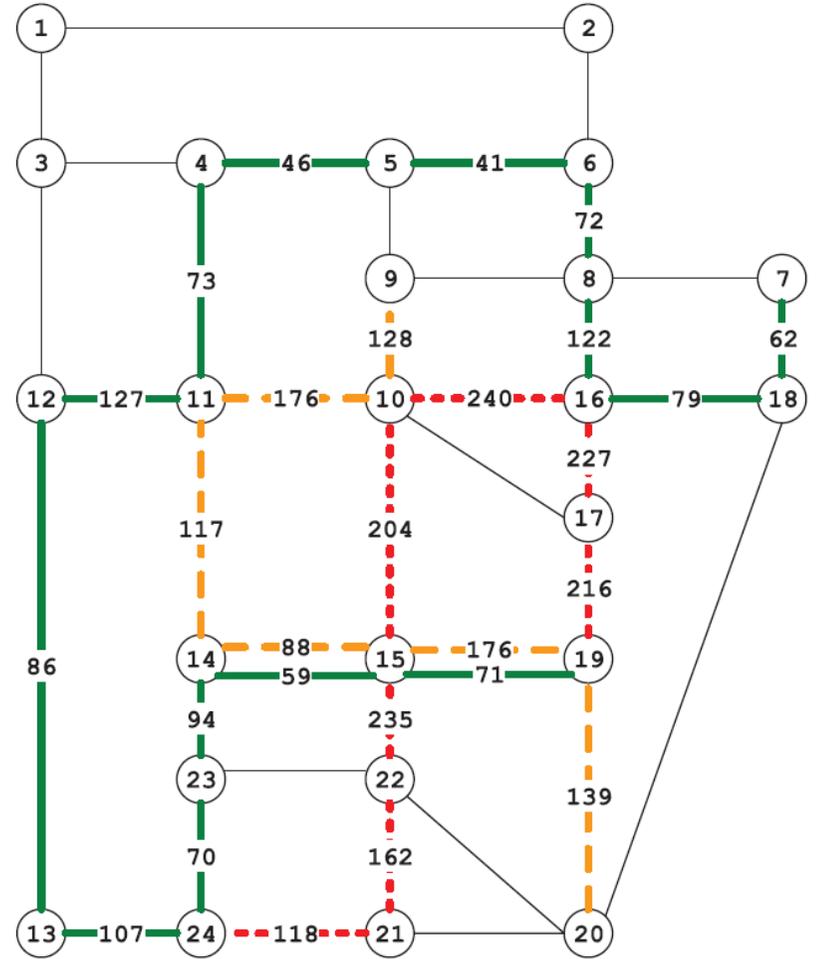
- The SR-based approach obtains the optimal level of robustness with lowest total cost
- Higher level of robustness indicates higher transit cost and lower Dial-a-ride cost

		Level of robustness	Company cost		Passenger cost		Total cost
			RTL	Dial-a-ride	RTL	Dial-a-ride	
SO	SR-based method	0.61	1111	107	861	22	2100
	$0^* \sigma$	0.50	1145	112	863	23	2143
	$1^* \sigma$	0.84	1234	23	885	5	2147
	$2^* \sigma$	0.97	1346	2	875	0	2223
	$3^* \sigma$	0.998	1455	1	739	0	2195
UE	SR-based method	0.67	1305	117	697	23	2142
	$0^* \sigma$	0.50	1145	170	852	34	2201
	$1^* \sigma$	0.84	1234	114	953	23	2324
	$2^* \sigma$	0.97	1346	86	916	17	2365
	$3^* \sigma$	0.998	1455	11	736	2	2204



- 114 Link flow/100
- - - Line 1[#] Freq.=12 TU/h
- - - Line 2[#] Freq.=9.2 TU/h
- Line 3[#] Freq.=8 TU/h

(a) Solutions under SO



- 94 Link flow/100
- - - Line 1[#] Freq.=15 TU/h
- - - Line 2[#] Freq.=11 TU/h
- Line 3[#] Freq.=8 TU/h

(b) Solutions under UE



Conclusions

- Schedule based ferry service network design problem under demand uncertainty
 - Combination of regular ferry services and ad-hoc services
 - Service reliability to separate the two service types deployment into two phases
 - Two passenger flow patterns SO and UE are investigated
 - Value of reservation to the company under stochastic demand
- Frequency based rapid transit network design problem
 - Multi-line design without predetermined origin and destination
 - Transfer costs are accounted for
 - Consider the problem from the perspective of robust optimization
 - Service reliability conveys the level of robustness
 - Improve the current SR based gradient solution algorithm



Key Takeaways

- Jointly optimize FRT and DRT as an integrated system can achieve substantial cost efficiency in addressing stochastic demand.
- The notion of Service Reliability (SR) offers an efficient way to reformulate and solve stochastic programs, which also allows the incorporation of three difficult extensions in TND: addressing stochastic demand, hard capacity constraints, and user equilibrium conditions.