A Column Generation Algorithm for Choice-Based Network Revenue Management

Gustavo Vulcano
Stern School of Business
New York University

Joint work with J. Miranda Bront and I. Méndez
(Computer Science Department, University of Buenos Aires, Argentina)

7th Annual INFORMS Revenue Management and Pricing Section Conference
Barcelona, Spain – June 2007

Outline

- What is choice-based RM? Why is it important?
- Overview of a choice-based RM approach
- Choice-based LP formulation
- Column generation subproblem
- Decomposition approximation algorithm
- Numerical experiments
- Conclusions

Why does choice behavior matter?

- Traditional RM models: Independent demand assumption

BUT

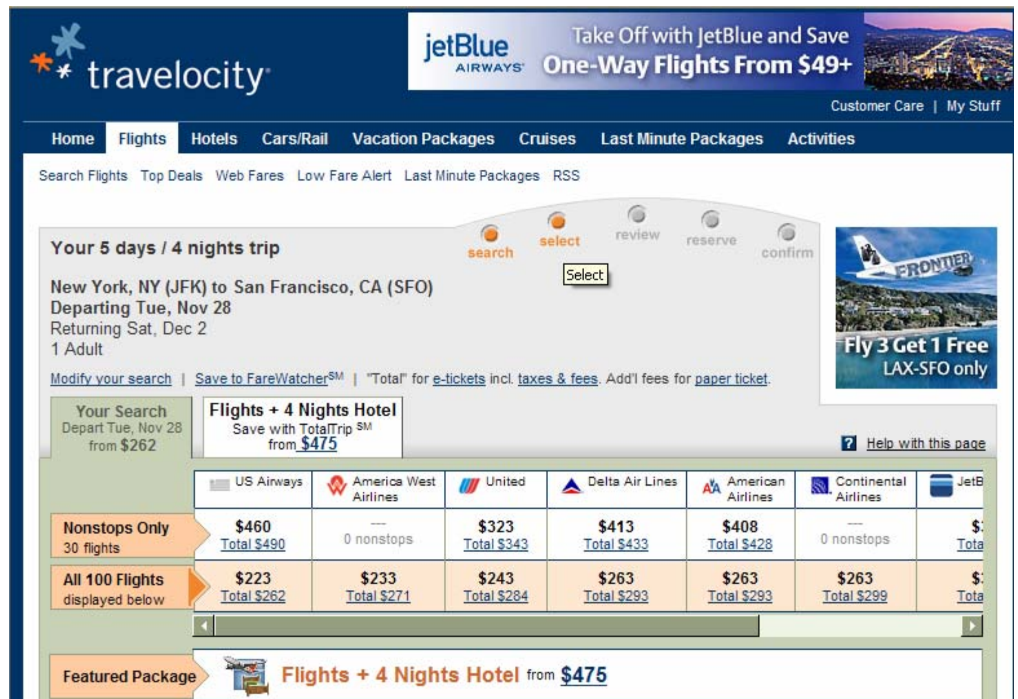
- Observed bookings are a function of product availability and customers' purchasing decisions

SO

- Likely that airline decisions could be dramatically improved by accounting explicitly for choice behavior effects

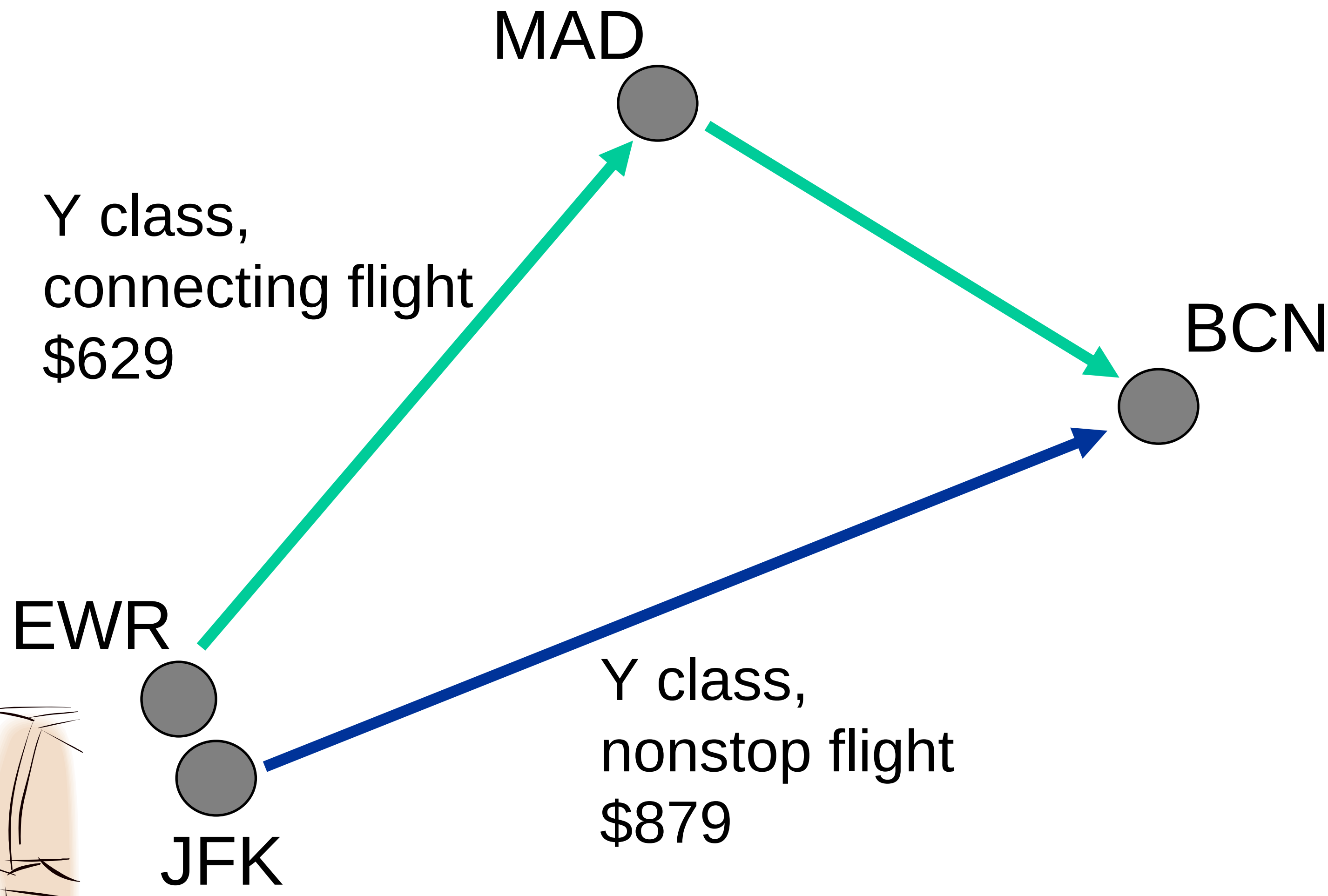
Demand for products is the outcome of a customer choice decision …

Customer chooses among the current set of alternatives



Wait, or not buy?

Choice among different alternatives affects revenue



What does it take to implement choice-based RM?

1. Models of customer choice behavior

2. Techniques for estimating choice models from available data

3. Revenue optimization methods that can deal with complex, choice-based models of demand

1. Modeling choice behavior

- A wealth of scientific work exists on how customers make choices
  - Classical utility theory
  - Behavioral economics
  - Marketing science

- No need to “reinvent the wheel” in this area; the question is how to <u>apply</u> this theory to RM

- Common practice: *utility maximization model* (e.g., Multinomial Logit model (MNL))

2. Estimating choice models from data

How do we estimate choice behavior using real-world airline data?

Two components to consider:

1) Choice behavior estimation

How will individual customers choose based on the available alternatives and their attributes?

Example: MNL model

2) Volume estimation

How many customers are making choices each day?

Example: Market share information, EM method

Choice behavior estimation

- Assume customers decide to purchase a flight based on attributes such as
  - Day of departure
  - Time of departure
  - Price
  - Departure airport
  - Airline brand

- The MNL is a parametric model that could describe and allow to estimate choice based on these attributes

Choice behavior estimation: Multinomial Logit (MNL) model

- Probability customer n purchases flight i

$$P_n(i) = \frac{e^{\beta^\top x_{in}}}{\sum_{j \in C_n^a} e^{\beta^\top x_{jn}} + 1}, \quad i \in C_n$$

where

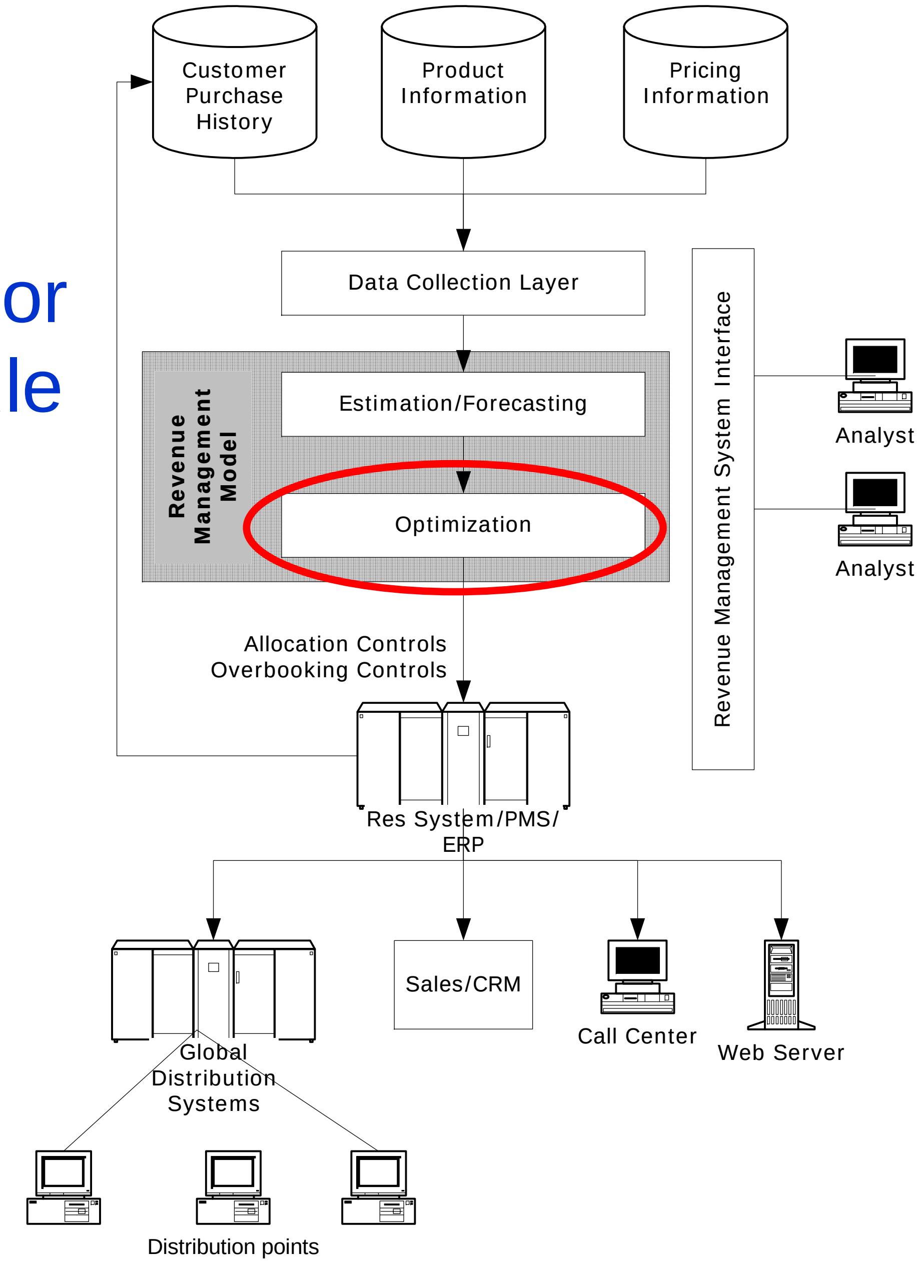
x_{jn} vector of observable attributes for alternative j available to customer n at time of purchase

β vector of weights (to be computed from data)

C_n choice set faced by customer n (major design decision; in our case, “same day flights”)

3. Revenue optimization

How can we embed the estimated choice-behavior in the optimization module of a RM system?



References

- Talluri & van Ryzin, The theory and practice of Revenue Management, Kluwer Academic Publishers, 2004.
- Talluri & van Ryzin, “RM under a general discrete choice model,” Management Science, Vol. 50, No. 1, 2004.
- Gallego, Iyengar, Phillips and Dubey. “Managing flexible products on a network”, Columbia University, School of Engineering and Applied Sciences. Working Paper, 2004.
- Zhang & Cooper “RM for parallel flights with customer choice behavior”, *Operations Research*, Vol. 53, pp.414-431, 2005.
- van Ryzin, G. J. “Models of demand”, *J. of Revenue and Pricing Mgmt.,* Vol. 4 No 2, 2005.
- van Ryzin & Liu, “On the choice-based linear programming model for network RM”, Graduate School of Business, Columbia University. Working Paper, 2004.
- Zhang and Adelman, “An approximate dynamic programming approach to network revenue management with customer choice”. University of Chicago. Working paper, 2006.
- Vulcano, G., van Ryzin, G. J. and Chaar, W., “Choice-Based Revenue Management: An Empirical Study of Estimation and Optimization,” in preparation.

Problem formulation

- Set of products $N = \{1,2,\ldots, n\}$, m resources
- Network described by incidence matrix $[A_{ij}]$, with

$$A_{ij} = \begin{cases} 1, & \text{if product } j \text{ uses resource } i \\ 0, & \text{otherwise} \end{cases}$$

- Remaining capacities $x = (x_1,\ldots, x_m)$, with initial capacities c
- Revenues $r = (r_1,\ldots, r_n)$
- Discrete time horizon $t=1,\ldots,T$, with probability λ of having one arrival in every small period
- Problem: Which subset of products $S \subseteq N$ to offer in every period t?
- DP formulation:

$$V_t(x) = \max_{S \subseteq N} \left\{ \sum_{j \in S} \lambda P_j(S)(r_j + V_{t+1}(x - A_j)) + (\lambda P_0(S) + 1 - \lambda) V_{t+1}(x) \right\}$$

$$= \max_{S \subseteq N} \left\{ \sum_{j \in S} \lambda P_j(S)(r_j - (V_{t+1}(x) - V_{t+1}(x - A_j))) \right\} + V_{t+1}(x)$$

$$\text{Boundary conditions}: V_t(0) = 0, \forall t; \quad V_{T+1}(x) = 0, \forall x$$

Choice-based LP (Gallego et. al. (2004))

Define:

- $R(S)$: expected revenue generated from an arrival when set S is offered; i.e. $R(S) = \sum_{j \in S} r_j P_j(S)$
- $Q_i(S)$: probability of using a unit of capacity on leg i when S is offered and a customer arrives; i.e. $Q(S) = AP(S)$
- $t(S)$: number of periods to offer set S (decision variables)

$$V^{CDLP} = \max \sum_{S \subseteq N} \lambda R(S) t(S) \qquad \text{[Revenue]}$$

$$\text{s.t.} \sum_{S \subseteq N} \lambda Q(S) t(S) \le c \qquad \text{[Capacity availability, dual vble. } \pi \text{]}$$

$$\sum_{S \subseteq N} t(S) \le T \qquad \text{[Time availability, dual vble. } \sigma \text{]}$$

$$t(S) \ge 0, \forall S \subseteq N$$

Properties of the CDLP

- Asymptotically optimal (van Ryzin & Liu (2004))
- Problem: Exponential number of variables
  Solution: Use column generation

Generic Column Generation Procedure

1. Select a restricted set of variables
2. Optimize the problem considering only this restricted set
3. Check if there exists a variable with positive reduced cost (COLGEN)
   If so, add it to the set and go to Step 1.

- In our case, COLGEN is given by:

$$\max_{S \subseteq N} \left\{ \lambda R(S) - \lambda \pi^{\mathsf{T}} Q(S) \right\} - \sigma$$

Model of Market Segmentation

- Customers belong to different segments $l = 1,\ldots, L$

- Each segment is defined by one consideration set $C_l \subseteq N$. A customer belongs to segment l with probability p_l, so that $\lambda_l = \lambda\, p_l$, and

$$P_{lj}(S) = \frac{v_{lj}}{\sum_{h \in C_l \bigcap S} v_{lh} + v_{l0}}$$

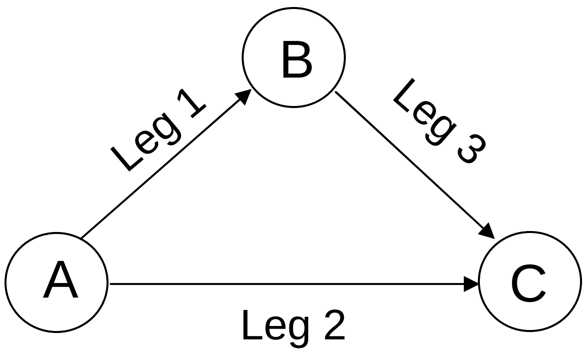
The probability that the airline sells a unit of product j when offering set S, conditional on having an arrival, is :

$$P_j(S) = \sum_{l=1}^{L} p_l\, P_{lj}(S)$$

- Limitations of the existing literature:
  - Gallego et al. (2004): No market segmentation allowed, i.e. $C_l = N$.
  - van Ryzin and Liu (2004): Disjoint segments, i.e. $C_l \cap C_h = \emptyset$ for $l \neq h$.
  - Zhang and Adelman (2006): Approx. DP, disjoint segments

Simple example

Network



Product definitions

Product	O-D	Class	Fare
1	A, C	H	1200
2	A, B, C	H	800
3	A, B	H	500
4	B, C	H	500
5	A, C	L	800
6	A, B, C	L	500
7	A, B	L	300
8	B, C	L	300

Segment definitions

Segment	λ_l	Consideration set	Pref. vector	Description
1	0.15	{1, 5}	(5, 8, 2)	Price sensit., nonstop A $\rightarrow$ C
2	0.15	{1, 2}	(10, 6, 5)	Price insensit., A $\rightarrow$ C
3	0.20	{5, 6}	(8, 5, 2)	Price sensit., A $\rightarrow$ C
4	0.25	{3, 7}	(4, 8, 2)	Price sensit., A $\rightarrow$ B
5	0.25	{4, 8}	(6, 8, 2)	Price sensit., B $\rightarrow$ C

Column generation subproblem

- For our overlapping-segment case under the MNL model, the COLGEN problem

$$\max_{S \subset N} \left\{ \lambda R(S) - \lambda \pi^{\top} Q(S) \right\} - \sigma$$

becomes

$$\max_{y \in \{0,1\}^n} \left\{ \sum_{j=1}^{n} \left(r_j - A_j^{\top} \pi \right) y_j \left(\sum_{l=1}^{L} \frac{\lambda_l v_{lj}}{\sum_{i \in C_l} v_{li} y_i + v_{l0}} \right) \right\} - \sigma$$

- If its optimal function value is non-positive, then π and σ are dual feasible, and the current solution of the CDLP is optimal.

Complexity of COLGEN

- It is a particular case of the NP-Hard, hyperbolic binary programming problem:

$$\max_{y\in\{0,1\}^n} \sum_{l=1}^{L} \frac{a_{l0} + \sum_{j=1}^{n} a_{lj} y_j}{b_{l0} + \sum_{j=1}^{n} b_{lj} y_j}$$

where

$$a_{l0} + \sum_{j=1}^{n} a_{lj} y_j \geq 0 \quad \text{and} \quad b_{l0} + \sum_{j=1}^{n} b_{lj} y_j > 0$$

- Theorem: The hyperbolic binary programming problem

$$\max_{y\in\{0,1\}^n} \sum_{j=1}^{n} w_j y_j \left(\sum_{l=1}^{L} \frac{\lambda_l v_{lj}}{\sum_{i\in C_l} v_{li} y_i + v_{l0}} \right)$$

where w_j, $v_{lj}>0$, $l=1,\ldots, L$, $j=1,\ldots, n$, and $C_l \subseteq N$, is NP-Hard.

Proof: Polynomial transformation from *minimum vertex cover*.

Solving COLGEN

- Approach 1: Greedy Heuristic

  Start from an empty set S, and add sequentially the element that provides the maximum marginal increase to the current solution.

- Approach 2: MIP (exact) formulation (Prokopyev et. al. (2005)), based on a linearization procedure proposed by Wu (1997).

- Our implementation:
  - First try the Greedy Heuristic. If no entering column is found, then try the MIP formulation.

Decomposition approximation algorithm

- Phase 1: Assessing the value of capacity

  Approximate the network value function at leg i by

$$V_t(x) \approx \hat{V}_t^i(x_i) + \sum_{k \neq i} \pi_k^* x_k,$$

  where $\hat{V}_t^i(x_i)$ is a dynamic (time dependent) approx. of the value of capacity of leg i, and $\pi_k^* x_k$ are static (time indep.) approx. of the value of capacity of other legs

  Averaging out these approximations, we get

$$V_t(x) \approx \frac{1}{m} \sum_{i=1}^{m} \left(\hat{V}_t^i(x_i) + \sum_{k \neq i} \pi_k^* x_k \right)$$

Decomposition approximation algorithm (cont'd)

- How difficult is the single leg problem?

$$\hat{V}_t^i(x_i) = \max_{S \subset N} \left\{ \sum_{j \in S} \lambda P_j(S) \left(r_j - (\Delta \hat{V}_{t+1}^i(x_i) - \pi_i^*) \mathbb{1}\{i \in A_j\} - \sum_{k \in A_j} \pi_k^* \right) \right\} + \hat{V}_{t+1}^i(x_i)$$

- The single leg RM problem under choice behavior was studied by Talluri and van Ryzin (2004). They characterize sufficient and necessary conditions for the strategy "nesting by fare order" (sequence of nested offer sets, where the incremental elements follow the fare order) to be optimal

- Proposition: The MNL choice model with overlapping segments over a single leg does not admit "nesting by fare order" as optimal strategy.

- In fact, it shares the structure of COLGEN (i.e., it is NP-Hard)

Decomposition approximation algorithm (cont'd)

- Phase 2: Computing the offer sets dynamically

  Marginal value of capacity:

$$\Delta V_t^i(x) \approx \frac{1}{m}\Delta \hat{V}_t^i(x_i) + \frac{m-1}{m}\pi_i^*$$

  Or more generally,

$$\Delta V_t^i(x) \approx \Delta \bar{V}_t^i(x) := \beta \Delta \hat{V}_t^i(x_i) + (1-\beta)\pi_i^*, \quad \text{for } 0 \le \beta \le 1.$$

  The firm must select a set S dynamically in each period t, by solving

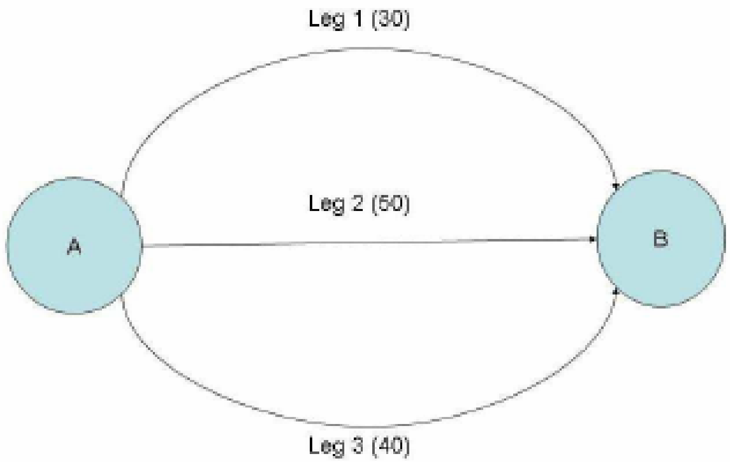
$$\max_{S \subset N} \left\{ \sum_{j \in S, j \text{ available}} \lambda P_j(S) \left(r_j - \Delta \bar{V}_t^{\mathsf{T}}(x) A_j \right) \right\} \qquad (*)$$

Numerical experiments

Methods tested

- DCOMP: Solve CDLP, and use β=1 to compute the marginal value of capacity
- DCOMP-0.5: Solve CDLP, and use β=0.5 to compute the marginal value of capacity
- CDLP: Use the primal outcome of CDLP, and offer the sets following the indices of the variables
- RCDLP: Randomized CDLP; shuffle the indices of the variables.
- ROPT-freq: Reoptimize CDLP with a predetermined frequency in order to get updated values of the dual variables.
- INDEP: Solve the DLP, for which the mean demand of product j is set at $\lambda\, P_j(N)\, T$ (i.e., all products are simultaneously offered).

Example 1: Parallel Flights



Airline network

Product	Leg	Class	Fare
1	1	L	400
2	1	H	800
3	2	L	500
4	2	H	1000
5	3	L	300
6	3	H	600

Product definition

Booking horizon: T=300 periods, with an average of 150 arrivals per stream

Segment	Consideration set	Pref. vector	λ_l	Description
1	{2,4,6}	(5,10,1)	0.1	Price insensitive, afternoon preference
2	{1,3,5}	(5,1,10)	0.15	Price sensitive, evening preference
3	{1,2,3,4,5,6}	(10,8,6,4,3,1)	0.2	Early preference, price sensitive
4	{1,2,3,4,5,6}	(8,10,4,6,1,3)	0.05	Price insensitive, early preference

Table 4: Segment definitions for Parallel Flights instance

Example 1: Revenue results

α	v_0	UB	DCOMP		DCOMP-0.5		CDLP		RCDLP		ROPT-0.01		ROPT-0.1		INDEP	
			Mean	%LF	Mean	%LF	Mean	%LF	Mean	%LF	Mean	%LF	Mean	%LF	Mean	%LF
0.6	(1,5,5,1)	56,884	55,948	98.60	55,872	97.82	54,177	94.60	53,783	94.42	51,632	95.83	49,468	93.48	49,794	97.41
	(1,10,5,1)	56,848	55,882	98.36	55,698	97.29	54,051	94.45	53,619	94.08	51,946	93.81	49,994	91.01	49,655	96.77
	(5,20,10,5)	53,819	51,326	94.59	51,338	94.23	50,068	92.94	50,476	93.44	48,602	97.35	46,274	95.05	46,246	92.53
0.8	(1,5,5,1)	71,936	69,533	95.78	69,163	94.79	68,105	94.69	68,641	95.11	66,336	94.57	64,490	92.75	60,346	94.14
	(1,10,5,1)	71,794	69,129	95.24	68,863	94.50	67,896	94.58	68,491	95.18	66,237	93.82	64,522	91.91	59,552	91.40
	(5,20,10,5)	61,868	60,147	90.48	60,222	90.49	59,073	90.82	59,289	90.83	56,724	91.98	55,842	91.16	53,044	81.94
1.0	(1,5,5,1)	79,155	76,954	95.65	77,096	95.91	75,726	94.96	75,996	94.89	77,106	95.36	76,136	94.36	66,224	85.38
	(1,10,5,1)	76,866	75,639	90.88	75,635	90.97	73,788	90.26	74,100	90.19	75,299	91.26	74,635	90.73	64,831	81.76
	(5,20,10,5)	63,255	62,775	78.09	62,792	78.15	62,702	78.41	62,541	78.41	62,040	79.94	61,827	79.96	56,203	72.44
1.2	(1,5,5,1)	80,371	79,817	84.28	79,818	84.24	79,666	84.33	79,698	84.32	79,834	84.58	79,774	84.53	68,970	76.84
	(1,10,5,1)	78,045	77,520	79.06	77,526	79.07	77,348	79.26	77,332	79.22	77,529	79.49	77,476	79.45	67,570	73.71
	(5,20,10,5)	63,296	63,111	67.52	63,113	67.52	62,491	68.88	62,677	68.87	62,422	70.20	62,293	70.20	58,543	65.88
1.4	(1,5,5,1)	81,066	80,408	73.08	80,376	72.83	80,362	72.75	80,362	72.75	80,439	73.27	80,421	73.22	71,418	70.60
	(1,10,5,1)	78,816	78,123	68.56	78,097	68.32	78,091	68.24	78,091	68.24	78,136	68.80	78,120	68.76	69,949	67.87
	(5,20,10,5)	63,237	63,211	60.54	63,212	60.54	62,553	62.04	62,775	62.00	62,822	62.82	62,734	62.85	60,732	60.82

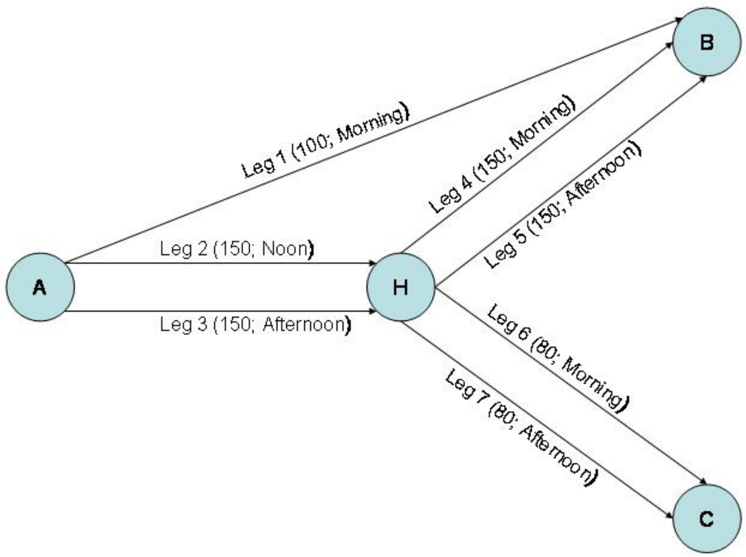
Table 5: Revenue results for the Parallel Flights example.

Example 1: Revenue gaps with respect to DCOMP

α	v_0	% DCOMP-0.5	% CDLP	% RCDLP	% ROPT-0.01	% ROPT-0.1	% INDEP
0.6	(1,5,5,1)	0.13 ± 0.18	3.16 ± 0.31	3.86 ± 0.31	7.71 ± 0.24	11.58 ± 0.28	10.90 ± 0.23
	(1,10,5,1)	0.32 ± 0.20	3.27 ± 0.32	4.04 ± 0.32	7.04 ± 0.27	10.53 ± 0.29	11.14 ± 0.23
	(5,20,10,5)	-0.02 ± 0.39	2.47 ± 0.41	1.65 ± 0.41	5.30 ± 0.36	9.84 ± 0.39	9.89 ± 0.37
0.8	(1,5,5,1)	0.53 ± 0.33	2.05 ± 0.33	1.28 ± 0.35	4.59 ± 0.30	7.25 ± 0.32	13.21 ± 0.33
	(1,10,5,1)	0.38 ± 0.35	1.78 ± 0.34	0.92 ± 0.36	4.18 ± 0.32	6.66 ± 0.33	13.85 ± 0.34
	(5,20,10,5)	-0.12 ± 0.47	1.78 ± 0.44	1.42 ± 0.44	5.69 ± 0.42	7.15 ± 0.41	11.80 ± 0.46
1.0	(1,5,5,1)	-0.18 ± 0.34	1.59 ± 0.33	1.24 ± 0.33	-0.19 ± 0.34	1.06 ± 0.34	13.94 ± 0.38
	(1,10,5,1)	-0.02 ± 0.42	2.44 ± 0.38	2.03 ± 0.39	0.31 ± 0.23	1.32 ± 0.40	14.28 ± 0.41
	(5,20,10,5)	-0.53 ± 0.56	0.11 ± 0.54	0.37 ± 0.54	1.17 ± 0.54	1.51 ± 0.54	10.46 ± 0.52
1.2	(1,5,5,1)	0.00 ± 0.46	0.18 ± 0.45	0.14 ± 0.45	-0.02 ± 0.29	0.05 ± 0.45	13.58 ± 0.43
	(1,10,5,1)	0.00 ± 0.48	0.22 ± 0.47	0.24 ± 0.47	-0.01 ± 0.48	0.05 ± 0.47	12.83 ± 0.45
	(5,20,10,5)	0.00 ± 0.57	0.98 ± 0.56	0.68 ± 0.56	1.09 ± 0.57	1.29 ± 0.63	7.23 ± 0.53
1.4	(1,5,5,1)	0.03 ± 0.48	0.05 ± 0.47	0.05 ± 0.47	-0.03 ± 0.49	-0.01 ± 0.49	11.18 ± 0.45
	(1,10,5,1)	0.03 ± 0.50	0.04 ± 0.50	0.04 ± 0.50	-0.01 ± 0.50	0.00 ± 0.50	10.46 ± 0.46
	(5,20,10,5)	0.00 ± 0.56	1.04 ± 0.56	0.68 ± 0.55	0.61 ± 0.56	0.75 ± 0.56	3.92 ± 0.54

Table 6: 95% confidence intervals for the suboptimality gaps with respect to DCOMP in the Parallel Flights example. Simulation results based on 2,000 streams of demand per scenario.

Example 2: Small Network



Product definitions

Product	Legs	Class	Fare	Product	Legs	Class	Fare
1	1	H	1000	12	1	L	500
2	2	H	400	13	2	L	200
3	3	H	400	14	3	L	200
4	4	H	300	15	4	L	150
5	5	H	300	16	5	L	150
6	6	H	500	17	6	L	250
7	7	H	500	18	7	L	250
8	{2,4}	H	600	19	{2,4}	L	300
9	{3,5}	H	600	20	{3,5}	L	300
10	{2,6}	H	700	21	{2,6}	L	350
11	{3,7}	H	700	22	{3,7}	L	350

Booking horizon: T=1,000 periods, , with an average of 910 arrivals per stream

Segment	O-D	Consideration set	Pref. vector	λ_l	Description
1	$A \to B$	{1,8,9,12,19,20}	(10,8,8,6,4,4)	0.08	Price insensitive, early pref.
2	$A \to B$	{1,8,9,12,19,20}	(1,2,2,8,10,10)	0.2	Price sensitive
3	$A \to H$	{2,3,13,14}	(10,10,5,5)	0.05	Price insensitive
4	$A \to H$	{2,3,13,14}	(2,2,10,10)	0.2	Price sensitive
5	$H \to B$	{4,5,15,16}	(10,10,5,5)	0.1	Price insensitive
6	$H \to B$	{4,5,15,16}	(2,2,10,8)	0.15	Price sensitive, slight early pref.
7	$H \to C$	{6,7,17,18}	(10,8,5,5)	0.02	Price insensitive, slight early pref.
8	$H \to C$	{6,7,17,18}	(2,2,10,8)	0.05	Price sensitive
9	$A \to C$	{10,11,21,22}	(10,8,5,5)	0.02	Price insensitive, slight early pref.
10	$A \to C$	{10,11,21,22}	(2,2,10,10)	0.04	Price sensitive

Segment definitions

Example 2: Revenue results

α	v_0	UB	DCOMP		DCOMP-0.5		CDLP		RCDLP		ROPT-0.01		ROPT-0.1		INDEP	
			Mean	%LF	Mean	%LF	Mean	%LF	Mean	%LF	Mean	%LF	Mean	%LF	Mean	%LF
0.6	(1,5)	215,793	197,038	88.90	196,920	88.06	207,890	91.27	208,476	91.70	200,444	93.37	195,291	92.09	172,362	97.71
	(5,10)	200,515	194,146	93.35	191,443	92.09	194,393	91.90	193,658	91.82	192,896	95.07	189,182	94.42	163,905	96.76
	(10,20)	170,137	167,866	92.68	167,902	92.79	164,089	91.45	164,296	91.39	166,919	93.07	165,516	92.47	151,801	92.48
0.8	(1,5)	266,934	262,823	86.79	263,023	86.37	261,264	85.62	260,820	85.68	262,013	86.90	249,221	86.21	204,572	94.60
	(5,10)	223,173	220,891	90.48	221,012	90.51	215,884	89.38	217,105	89.40	217,073	90.75	214,192	90.15	191,066	90.22
	(10,20)	188,574	186,219	85.29	185,969	85.27	184,182	84.86	184,289	84.92	186,325	85.61	185,841	85.26	172,246	84.09
1.0	(1,5)	281,967	279,506	81.34	279,536	81.36	277,738	80.80	277,473	80.78	278,344	81.04	275,016	80.53	226,002	87.71
	(5,10)	235,284	233,929	84.41	233,891	84.27	230,342	83.86	231,250	83.89	233,138	84.09	232,376	83.68	209,701	83.64
	(10,20)	192,038	191,646	76.10	191,623	76.05	190,283	76.34	190,393	76.34	191,727	75.89	191,627	75.70	188,058	76.73
1.2	(1,5)	284,772	284,736	71.85	284,747	71.85	282,842	71.55	282,996	71.51	283,280	72.47	281,926	72.46	243,930	82.48
	(5,10)	238,562	238,539	72.38	238,502	72.26	238,299	72.03	238,299	72.03	238,548	72.35	238,523	72.33	225,691	77.65
	(10,20)	192,373	192,530	65.86	192,524	65.87	192,511	65.88	192,511	65.88	192,532	65.87	192,526	65.87	192,416	65.80
1.4	(1,5)	287,076	286,743	62.14	286,629	62.16	285,417	61.96	285,598	61.95	286,160	62.24	285,783	62.21	259,039	76.96
	(5,10)	238,562	238,843	61.80	238,843	61.80	238,843	61.80	238,843	61.80	238,843	61.80	238,843	61.80	231,937	68.82
	(10,20)	192,373	192,541	56.48	192,541	56.48	192,541	56.48	192,541	56.48	192,541	56.48	192,541	56.48	192,468	56.42

Example 2: Revenue gaps with respect to DCOMP

α	v_0	% DCOMP-0.5	% CDLP	% RCDLP	% ROPT-0.01	% ROPT-0.1	% INDEP
0.6	(1,5)	0.05 $\pm$ 0.16	-5.50 $\pm$ 0.18	-5.80 $\pm$ 0.18	-1.72 $\pm$ 0.17	0.88 $\pm$ 0.17	12.52 $\pm$ 0.15
	(5,10)	1.39 $\pm$ 0.17	-0.12 $\pm$ 0.18	0.25 $\pm$ 0.18	0.64 $\pm$ 0.16	2.55 $\pm$ 0.17	15.57 $\pm$ 0.15
	(10,20)	-0.02 $\pm$ 0.18	2.25 $\pm$ 0.19	2.12 $\pm$ 0.18	0.56 $\pm$ 0.17	1.39 $\pm$ 0.17	9.57 $\pm$ 0.16
0.8	(1,5)	-0.07 $\pm$ 0.13	0.59 $\pm$ 0.15	0.76 $\pm$ 0.15	4.11 $\pm$ 0.14	5.17 $\pm$ 0.15	22.16 $\pm$ 0.13
	(5,10)	-0.05 $\pm$ 0.15	2.26 $\pm$ 0.15	1.71 $\pm$ 0.16	1.72 $\pm$ 0.15	3.03 $\pm$ 0.16	13.50 $\pm$ 0.16
	(10,20)	0.13 $\pm$ 0.20	1.09 $\pm$ 0.19	1.03 $\pm$ 0.19	-0.05 $\pm$ 0.20	0.20 $\pm$ 0.21	7.50 $\pm$ 0.20
1.0	(1,5)	-0.01 $\pm$ 0.14	0.63 $\pm$ 0.15	0.72 $\pm$ 0.16	0.41 $\pm$ 0.15	1.60 $\pm$ 0.15	19.14 $\pm$ 0.15
	(5,10)	0.01 $\pm$ 0.16	1.53 $\pm$ 0.16	1.10 $\pm$ 0.17	0.33 $\pm$ 0.19	0.66 $\pm$ 0.17	10.35 $\pm$ 0.17
	(10,20)	0.01 $\pm$ 0.22	0.71 $\pm$ 0.22	0.65 $\pm$ 0.22	-0.04 $\pm$ 0.30	0.00 $\pm$ 0.22	1.87 $\pm$ 0.22
1.2	(1,5)	0.00 $\pm$ 0.16	0.66 $\pm$ 0.16	0.61 $\pm$ 0.17	0.51 $\pm$ 0.17	0.98 $\pm$ 0.16	14.33 $\pm$ 0.16
	(5,10)	0.01 $\pm$ 0.19	0.10 $\pm$ 0.18	0.10 $\pm$ 0.18	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19	5.38 $\pm$ 0.18
	(10,20)	0.00 $\pm$ 0.22	0.00 $\pm$ 0.23	0.00 $\pm$ 0.23	0.00 $\pm$ 0.22	0.00 $\pm$ 0.22	0.05 $\pm$ 0.21
1.4	(1,5)	0.03 $\pm$ 0.17	0.46 $\pm$ 0.17	0.39 $\pm$ 0.16	0.20 $\pm$ 0.17	0.33 $\pm$ 0.17	9.66 $\pm$ 0.16
	(5,10)	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19	2.89 $\pm$ 0.19
	(10,20)	0.00 $\pm$ 0.22	0.00 $\pm$ 0.22	0.00 $\pm$ 0.22	0.00 $\pm$ 0.22	0.00 $\pm$ 0.22	0.03 $\pm$ 0.22

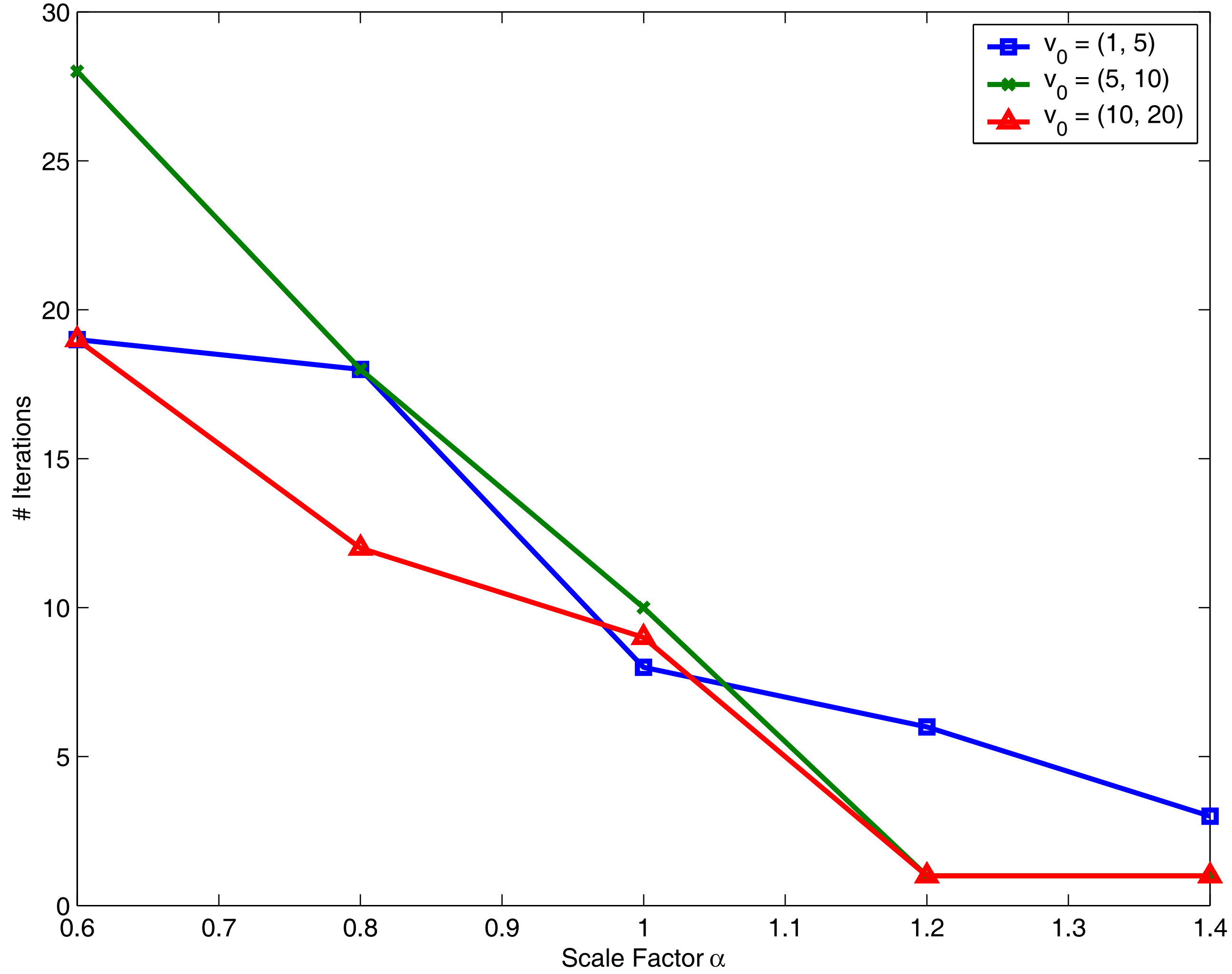
Simulation results based on 2,000 streams of demand per scenario.

Example 2: Computational times (in seconds)

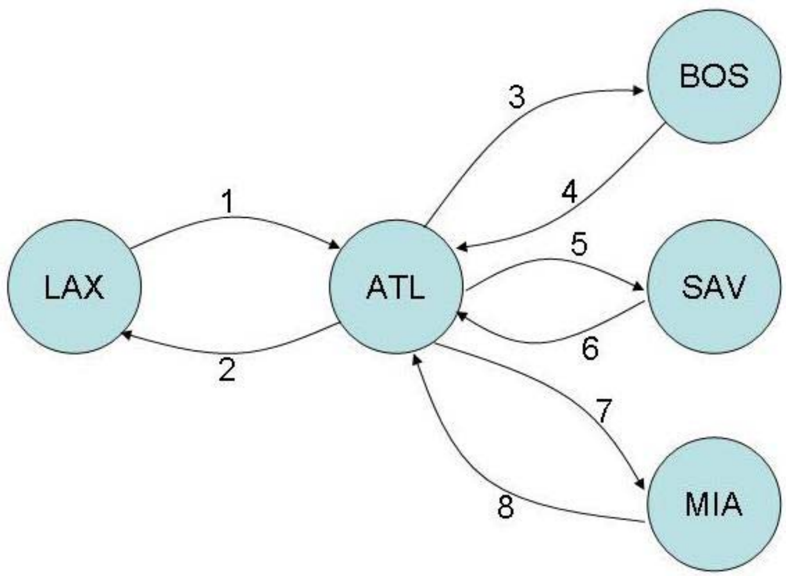
α	v_0	COLGEN	V_s^*	Solution of (*)	ROPT-0.01	ROPT-0.1	INDEP
0.6	(1,5)	0.05	171.81	0.01	6.64	0.75	0.30
	(5,10)	0.11	246.14	0.02	11.00	1.12	0.31
	(10,20)	0.07	282.52	0.02	6.08	0.64	0.31
0.8	(1,5)	0.07	285.43	0.03	7.71	0.83	0.42
	(5,10)	0.15	372.04	0.09	10.41	1.03	0.42
	(10,20)	0.05	406.98	0.02	4.19	0.41	0.43
1.0	(1,5)	0.09	409.9	0.06	7.43	0.78	0.54
	(5,10)	0.08	496.79	0.04	4.52	0.44	0.53
	(10,20)	0.05	516.30	0.02	2.74	0.27	0.54
1.2	(1,5)	0.09	506.72	0.06	7.09	0.74	0.64
	(5,10)	0.04	616.86	0.02	2.74	0.27	0.64
	(10,20)	0.04	636.43	0.01	2.29	0.23	0.67
1.4	(1,5)	0.07	599.47	0.05	6.53	0.67	0.75
	(5,10)	0.04	729.83	0.02	2.62	0.26	0.78
	(10,20)	0.03	737.17	0.01	2.29	0.22	0.79

Note: Exact MIP formulation used for solving (*)

Example 2: Iterations of COLGEN



Example 3: Hub-and-Spoke Network



O-D Market	Legs	Revenue			
		Y	M	B	Q
ATLBOS/BOSATL	3/4	310	290	95	69
ATLLAX/LAXATL	2/1	455	391	142	122
ATLMIA/MIAATL	7/8	280	209	94	59
ATLSAV/SAVATL	5/6	159	140	64	49
BOSLAX/LAXBOS	4-2/1-3	575	380	159	139
BOSMIA/MIABOS	4-7/8-3	403	314	124	89
BOSSAV/SAVBOS	4-5/6-3	319	250	109	69
LAXMIA/MIALAX	1-7/8-2	477	239	139	119
LAXSAV/SAVLAX	1-5/6-2	502	450	154	134
MIASAV/SAVMIA	8-5/6-7	226	168	84	59

Product definitions

Setting:

- 8 legs
- 80 products
- 40 customer segments
- Booking horizon: T =2,000, with an average of 1,732 arrivals per stream of demand.

Example 3: Segment definitions

Segment	C_l	v_l	λ_l	Segment	C_l	v_l	λ_l
ATL/BOS H	{1,2,3,4}	{6,7,9,10}	0.015	BOS/MIA H	{41,42,43,44}	{6,7,10,10}	0.008
ATL/BOS L	{3,4}	{8,10}	0.035	BOS/MIA L	{43,44}	{8,10}	0.03
BOS/ATL H	{5,6,7,8}	{6,7,9,10}	0.015	MIA/BOS H	{45,46,47,48}	{6,7,10,10}	0.008
BOS/ATL L	{7,8}	{8,10}	0.035	MIA/BOS L	{47,48}	{8,10}	0.03
ATL/LAX H	{9,10,11,12}	{5,6,9,10}	0.01	BOS/SAV H	{49,50,51,52}	{5,6,9,10}	0.01
ATL/LAX L	{11,12}	{10,10}	0.04	BOS/SAV L	{51,52}	{8,10}	0.035
LAX/ATL H	{13,14,15,16}	{5,6,9,10}	0.01	SAV/BOS H	{53,54,55,56}	{5,6,9,10}	0.01
LAX/ATL L	{15,16}	{10,10}	0.04	SAV/BOS L	{55,56}	{8,10}	0.035
ATL/MIA H	{17,18,19,20}	{5,5,10,10}	0.012	LAX/MIA H	{57,58,59,60}	{5,6,10,10}	0.012
ATL/MIA L	{19,20}	{8,10}	0.035	LAX/MIA L	{59,60}	{9,10}	0.028
MIA/ATL H	{21,22,23,24}	{5,5,10,10}	0.012	MIA/LAX H	{61,62,63,64}	{5,6,10,10}	0.012
MIA/ATL L	{23,24}	{8,10}	0.035	MIA/LAX L	{63,64}	{9,10}	0.028
ATL/SAV H	{25,26,27,28}	{4,5,8,9}	0.01	LAX/SAV H	{65,66,67,68}	{6,7,10,10}	0.016
ATL/SAV L	{27,28}	{7,10}	0.03	LAX/SAV L	{67,68}	{10,10}	0.03
SAV/ATL H	{29,30,31,32}	{4,5,8,9}	0.01	SAV/LAX H	{69,70,71,72}	{6,7,10,10}	0.016
SAV/ATL L	{31,32}	{7,10}	0.03	SAV/LAX L	{71,72}	{10,10}	0.03
BOS/LAX H	{33,34,35,36}	{5,5,7,10}	0.01	MIA/SAV H	{73,74,75,76}	{6,7,8,10}	0.01
BOS/LAX L	{35,36}	{9,10}	0.032	MIA/SAV L	{75,76}	{9,10}	0.025
LAX/BOS H	{37,38,39,40}	{5,5,7,10}	0.01	MIA/SAV H	{77,78,79,80}	{6,7,8,10}	0.01
LAX/BOS L	{39,40}	{9,10}	0.032	MIA/SAV L	{79,80}	{9,10}	0.025

Example 3: Revenue results

α	v_0	UB	DCOMP		DCOMP-0.5		CDLP		RCDLP		INDEP	
			Mean	%LF	Mean	%LF	Mean	%LF	Mean	%LF	Mean	%LF
0.6	(1,5)	163,897	160,624	97.10	160,206	95.03	156,557	95.70	156,410	95.72	110,471	98.64
	(5,10)	132,674	130,971	97.68	130,875	97.29	126,425	95.45	126,822	95.38	104,230	98.30
	(10,20)	111,897	110,314	97.61	110,209	96.93	106,688	95.53	106,879	95.45	96,661	97.85
0.8	(1,5)	177,384	175,598	97.70	173,520	93.66	170,301	96.05	170,562	96.13	130,841	98.72
	(5,10)	146,338	144,597	97.44	144,377	96.99	140,857	95.90	140,671	95.93	123,209	98.37
	(10,20)	122,464	121,082	96.24	120,985	96.14	117,621	96.03	117,654	96.07	114,012	97.53
1.0	(1,5)	187,270	185,384	96.43	184,785	95.99	181,673	95.57	181,751	95.60	149,246	98.63
	(5,10)	156,243	154,718	94.52	154,508	94.21	151,907	95.03	151,832	95.02	140,161	98.03
	(10,20)	128,386	127,343	91.65	127,255	91.88	125,811	92.27	125,883	92.27	126,091	92.09
1.2	(1,5)	195,269	193,511	94.88	192,953	94.89	190,000	93.71	190,248	93.70	165,890	98.29
	(5,10)	160,206	159,386	87.28	159,354	87.37	157,877	87.36	157,922	87.35	154,210	95.57
	(10,20)	128,448	128,336	78.36	128,336	78.36	128,336	78.36	128,336	78.36	128,361	78.38
1.4	(1,5)	197,113	196,886	86.70	196,860	86.77	196,639	86.79	196,639	86.79	179,983	98.54
	(5,10)	160,453	160,350	76.28	160,352	76.28	160,350	76.28	160,350	76.28	159,435	88.54
	(10,20)	128,448	128,336	68.22	128,336	68.22	128,336	68.22	128,336	68.22	128,363	68.24

Example 3: Revenue gaps with respect to DCOMP

α	v_0	% DCOMP-0.5	% CDLP	% RCDLP	% INDEP
0.6	(1,5)	0.26 ± 0.42	2.53 ± 0.38	2.62 ± 0.37	31.22 ± 0.33
	(5,10)	0.07 ± 0.41	3.47 ± 0.40	3.16 ± 0.39	20.34 ± 0.35
	(10,20)	0.09 ± 0.41	3.28 ± 0.40	3.11 ± 0.41	12.37 ± 0.37
0.8	(1,5)	1.18 ± 0.38	3.01 ± 0.36	2.86 ± 0.35	25.48 ± 0.31
	(5,10)	0.15 ± 0.37	2.58 ± 0.36	2.71 ± 0.36	14.66 ± 0.31
	(10,20)	0.06 ± 0.37	2.84 ± 0.36	2.81 ± 0.36	5.82 ± 0.33
1.0	(1,5)	0.32 ± 0.33	2.00 ± 0.32	1.96 ± 0.32	19.49 ± 0.27
	(5,10)	0.13 ± 0.34	1.81 ± 0.32	1.86 ± 0.32	9.40 ± 0.29
	(10,20)	0.06 ± 0.36	1.20 ± 0.35	1.14 ± 0.35	0.98 ± 0.35
1.2	(1,5)	0.28 ± 0.30	1.81 ± 0.29	1.68 ± 0.29	14.27 ± 0.25
	(5,10)	0.02 ± 0.32	0.94 ± 0.31	0.91 ± 0.31	3.24 ± 0.28
	(10,20)	0.0 ± 0.39	0.0 ± 0.39	0.0 ± 0.39	-0.01 ± 0.40
1.4	(1,5)	0.00 ± 0.31	0.11 ± 0.31	0.11 ± 0.31	8.57 ± 0.27
	(5,10)	0.00 ± 0.35	0.00 ± 0.35	0.00 ± 0.35	0.57 ± 0.32
	(10,20)	0.00 ± 0.39	0.00 ± 0.39	0.00 ± 0.39	-0.02 ± 0.40

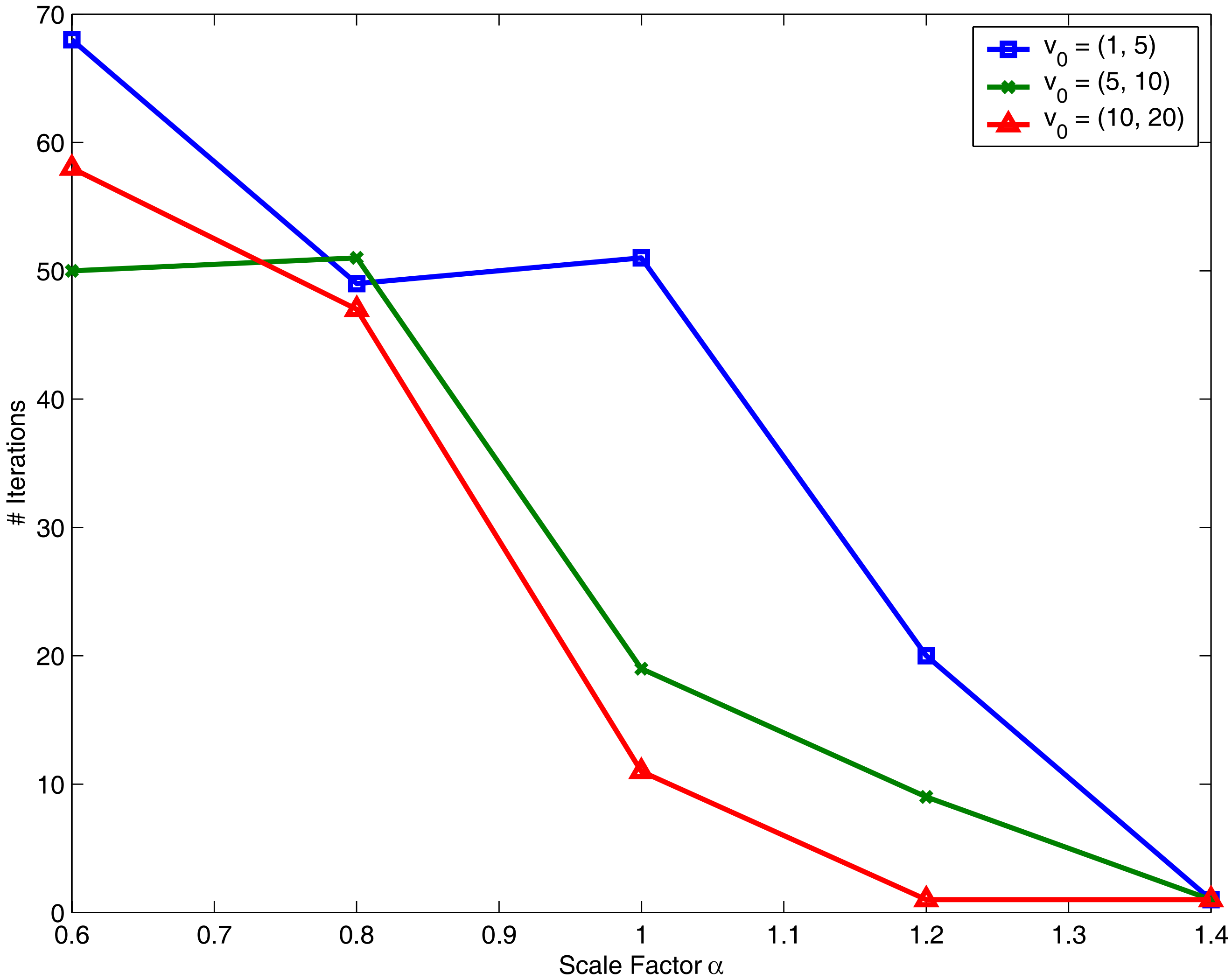
Simulation results based on 1,000 streams of demand per scenario.

Example 3: Computational times

α	v_0	COLGEN (seconds)	V_1^t (minutes)	Solution of (*) (seconds)	INDEP (seconds)
0.6	(1,5)	73.31	168.28	0.005	3.62
	(5,10)	0.70	205.11	0.006	3.66
	(10,20)	1.53	226.25	0.007	3.73
0.8	(1,5)	851.36	245.9	0.006	4.93
	(5,10)	5.75	295.40	0.007	4.99
	(10,20)	10.60	327.94	0.008	5.10
1.0	(1,5)	3184.22	341.30	0.006	6.31
	(5,10)	3.01	393.24	0.008	6.40
	(10,20)	0.25	422.22	0.008	6.39
1.2	(1,5)	1188.89	442.15	0.007	7.67
	(5,10)	10.55	490.50	0.007	7.67
	(10,20)	0.13	516.08	0.008	7.68
1.4	(1,5)	107.52	548.21	0.007	8.94
	(5,10)	0.85	576.21	0.007	8.96
	(10,20)	0.13	584.09	0.007	8.85

Note: Greedy Heuristic used for solving (*)

Example 3: Iterations of COLGEN



Conclusions

- Generalization of CDLP to the multiple, overlapping-segment case
- Need to develop a column generation algorithm to solve CDLP
- Even the column generation subproblem (COLGEN) is shown to be NP-Hard
- COLGEN: Simple, greedy heuristic is very effective to solve it
- Decomposition approximation (DCOMP) scheme of van Ryzin & Liu (2004) outperforms the other methods in terms of revenue
- Computational times of CDLP, RCDLP and ROPT are an order of magnitude shorter than DCOMP, while still leading to high quality results, conforming promising approached to implement in real-sized networks
- All choice-based methods tested outperform the INDEP model by more than 10%, suggesting that **choice behavior is a first order effect**