Reading Project – Cercetări Operaționale

Planificarea liniilor de transport public cu niveluri de serviciu prin *Column Generation*

Năstase Baraș Luca   Lupu Andreea   Dragoș Bădărău

14 ianuarie 2026

Cuprins

- Context și model: mulțimi, variabile, costuri, nivel de serviciu
- Dimensiune mare $\Rightarrow$ motivarea *Column Generation*
- RMP și dual: cost redus (legătura cu simplex/Lagrangian)
- Pricing: *shortest path cu resurse* și graf de stări (varianta clasică)
- Algoritmul CG + euristici uzuale
- Justificare: optimalitate LP, fracționaritate, *branch-and-price*

Notare (glosar rapid)

$G = (V, E)$: graful rețelei (noduri V, arce orientate E). Pentru $e = (u, v) \in E$, $t_e > 0$ este timpul pe arc.

$\mathcal{K}$: mulțimea cererilor OD. Pentru $k \in \mathcal{K}$:

$$(o_k, d_k, q_k, T_k^{\max}, \tau_k^{\max}),$$

unde o_k=origine, d_k=destinație, q_k=volum cerere, $T_k^{\max}$=timp maxim, $\tau_k^{\max}$=maxim transferuri.

$\mathcal{L}$: mulțimea liniilor fezabile; x_ℓ=frecvență, c_ℓ=cost operare.

$\mathcal{P}_k$: mulțimea rutelor fezabile pentru cererea k; $\mathcal{P} = \bigcup_{k \in \mathcal{K}} \mathcal{P}_k$.

$a_{\ell p} \in \{0, 1\}$: ruta p utilizează linia ℓ.

Context (rețea și cereri OD)

Rețeaua de transport: graf orientat $G = (V, E)$.
Cereri OD: $\mathcal{K}$, fiecare cerere $k \in \mathcal{K}$ are:

$$(o_k, d_k, q_k, T_k^{\max}, \tau_k^{\max}), \qquad q_k > 0.$$

Interpretare nivel de serviciu:

- $T_k^{\max}$ limitează *timpul total* al rutei;
- $\tau_k^{\max}$ limitează *numărul de transferuri*.

Mulțimi (linii și rute)

$\mathcal{L}$: mulțimea liniilor/serviciilor fezabile (fiecare linie este un traseu operat).
Pentru fiecare cerere $k \in \mathcal{K}$, $\mathcal{P}_k$: mulțimea rutelor fezabile (succesiuni de linii/arce) între o_k și d_k.
$\mathcal{P} = \bigcup_{k \in \mathcal{K}} \mathcal{P}_k$.

Indicatori de utilizare

Fie $a_{\ell p} \in \{0, 1\}$ indicatorul că linia $\ell \in \mathcal{L}$ este utilizată în ruta $p \in \mathcal{P}_k$.

Variabile decizionale (ce reprezintă fiecare)

Frecvențe de operare (decizia de planificare a liniilor)

Pentru fiecare linie $\ell \in \mathcal{L}$:

$$x_\ell \geq 0.$$

Semnificație: x_ℓ este frecvența/intenstitatea cu care operăm linia ℓ (de ex. vehicule/oră).

Alocarea cererii pe rute (decizia de rutare pasageri)

Pentru fiecare cerere $k \in \mathcal{K}$ și rută $p \in \mathcal{P}_k$:

$$y_{k,p} \in [0,1], \qquad \sum_{p \in \mathcal{P}_k} y_{k,p} = 1.$$

Semnificație: $y_{k,p}$ este fracția din cererea q_k trimisă pe ruta p (permite alocare fracționară în LP).

Costuri și mărimi de serviciu (definiții)

Cost operare linie: $c_\ell > 0$, $\forall \ell \in \mathcal{L}$.
Pentru rută $p \in \mathcal{P}_k$, definim:

$$T_{k,p} \text{ (timp total)}, \qquad \tau_{k,p} \text{ (nr. transferuri)}.$$

Cost generalizat pe unitatea de cerere

$$w_{k,p} = \alpha\, T_{k,p} + \beta\, \tau_{k,p}, \qquad \alpha, \beta > 0.$$

Semnificație: $w_{k,p}$ combină timpul și transferurile într-un singur cost perceput (ponderi α, β).

Fezabilitate (service level)

$$T_{k,p} \le T_k^{\max}, \qquad \tau_{k,p} \le \tau_k^{\max}.$$

Modelul primal (LP complet)

Obiectiv

$$\min \sum_{\ell \in \mathcal{L}} c_\ell x_\ell + \sum_{k \in \mathcal{K}} \sum_{p \in \mathcal{P}_k} w_{k,p} q_k \, y_{k,p}.$$

Constrângeri

$$\sum_{p \in \mathcal{P}_k} y_{k,p} = 1, \qquad \forall k \in \mathcal{K} \qquad (\mathrm{ce}$$

$$y_{k,p} \leq x_\ell, \qquad \forall k, \ \forall p \in \mathcal{P}_k, \ \forall$$

$$T_{k,p} \leq T_k^{\max}, \quad \tau_{k,p} \leq \tau_k^{\max}, \qquad \forall k, \ \forall p \in \mathcal{P}_k \qquad \text{(service level)}$$

$$x_\ell \geq 0, \quad 0 \leq y_{k,p} \leq 1.$$

De ce *Column Generation*?

$|\mathcal{P}_k|$ poate fi enormă (adesea exponențială) $\Rightarrow$ LP-ul complet este impracticabil.
Ideea: lucrăm cu un subset finit de rute $\mathcal{P}_k^t \subset \mathcal{P}_k$ și îl extindem iterativ.

Principiu

Rezolvăm un *Restricted Master Problem* (RMP) și căutăm rute cu *cost redus* negativ (pricing). Dacă nu există, soluția RMP este optimă pentru LP-ul complet.

Restricted Master Problem (RMP)

La iterația t, considerăm mulțimi finite:

$$\mathcal{L}^t \subset \mathcal{L}, \qquad \mathcal{P}_k^t \subset \mathcal{P}_k.$$

RMP

$$\min \sum_{\ell \in \mathcal{L}^t} c_\ell x_\ell + \sum_{k \in \mathcal{K}} \sum_{p \in \mathcal{P}_k^t} w_{k,p} q_k \, y_{k,p},$$

sub aceleași constrângeri, dar doar pentru $p \in \mathcal{P}_k^t$ și $\ell \in \mathcal{L}^t$.

Observație (implementare minimală)

În practică, $\mathcal{P}_k^t$ este menținut ca un *pool* (listă) de rute pentru fiecare cerere k.

Dual și cost redus: legătura cu simplex/Lagrangian

RMP este un LP; metoda simplex furnizează variabile duale pentru constrângeri. Costul redus al unei coloane $y_{k,p}$ are forma generală:

$$\bar{c}_{k,p} = c_{k,p} - \sum_i (\mathrm{dual}_i \cdot \mathrm{coef}_{i,k,p}), \qquad c_{k,p} = w_{k,p} q_k.$$

Echivalența cu formula simplex

$$\bar{c}_j = c_j - \pi^\top A_j.$$

O coloană poate îmbunătăți soluția curentă dacă $\bar{c}_{k,p} < 0$.

Cost redus pentru modelul cu activare $y \leq x$

Dualele relevante:

$\pi_k \in \mathbb{R}$ asociat constrângerii $\sum_{p \in \mathcal{P}_k^t} y_{k,p} = 1$;

$\mu_{k,p,\ell} \geq 0$ asociat inegalității $y_{k,p} - x_\ell \leq 0$ pentru fiecare ℓ cu $a_{\ell p} = 1$.

Formula completă (coloana $y_{k,p}$)

$$\boxed{\bar{c}_{k,p} = w_{k,p} q_k - \pi_k + \sum_{\ell \in \mathcal{L}^t:\ a_{\ell p}=1} \mu_{k,p,\ell}.}$$

Criteriu de generare

Dacă există k, p cu $\bar{c}_{k,p} < 0$, adăugăm ruta p în $\mathcal{P}_k^{t+1}$.

Problema de pricing (definiție)

Pentru fiecare cerere $k \in \mathcal{K}$, *pricing* caută ruta cu cost redus minim:

$$\min_{p \in \mathcal{P}_k} \ \bar{c}_{k,p} \quad \text{s.t.} \quad T_{k,p} \le T_k^{\max}, \ \tau_{k,p} \le \tau_k^{\max}.$$

Interpretare

Pricing este subproblema simplex: găsește coloane cu $\bar{c} < 0$.
Este o problemă de tip *Shortest Path Problem with Resource Constraints* (SPPRC).

De ce apare un *shortest path*? (structură aditivă)

Dacă rutele sunt secvențe de arce $p = (e_1, \ldots, e_m)$, atunci:

$$T_{k,p} = \sum_{j=1}^{m} t_{e_j}.$$

Cost pe rută (formă aditivă + termeni globali)

$$w_{k,p} q_k = \underbrace{\sum_{e \in p} q_k \alpha\, t_e}_{\text{aditiv pe arce}} + \underbrace{q_k \beta\, \tau_{k,p}}_{\text{depinde de transferuri}} .$$

Termenii aditivi pot fi integrați în costurile arcelor; constrângerile pe T și τ sunt resurse.

SPPRC clasic: graf de stări și semnificația variabilelor

În SPPRC clasic, o **stare** (label) este:

$$\boxed{(v,\ T,\ \tau)}$$

cu semnificație precisă:

- $v \in V$: nodul (stația) la care am ajuns;
- $T \in \mathbb{R}_{\geq 0}$: timpul acumulat pe drumul parcurs până la v;
- $\tau \in \mathbb{Z}_{\geq 0}$: numărul de transferuri acumulate până la v.

Fezabilitate în starea curentă

$$T \leq T_k^{\max}, \qquad \tau \leq \tau_k^{\max}.$$

Tranziții în graful de stări (cum se actualizează T și τ)

Pentru un arc fizic $e = (v, v') \in E$, o tranziție este:

$$(v, T, \tau) \longrightarrow (v', T + t_{(v,v')}, \tau + \Delta\tau),$$

unde:

- $T' = T + t_{(v,v')}$ acumulează timpul pe arc;
- $\Delta\tau$ este incrementul de transfer (depinzând de definiția transferului).

Cum definim $\Delta\tau$ (două variante standard)

- **Simplificat (ca în multe modele):** fiecare arc după primul induce un "transfer": $\Delta\tau = 1$ după primul pas.
- **Clasic în transport:** transferul se produce când se schimbă linia; atunci starea se extinde cu linia curentă (v, T, τ, ℓ) și $\Delta\tau = \mathbb{1}[\ell' \neq \ell]$.

Cost incremental în SPPRC (ce se pune pe arce)

Într-un algoritm de etichetare, un label păstrează și un cost acumulat. Partea aditivă pe arce este:

$$C \leftarrow C + q_k \alpha \, t_{(v,v')}.$$

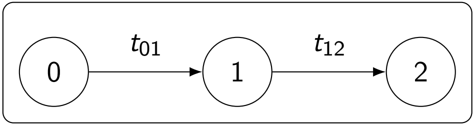
Termeni care nu sunt arc-cu-arc

$-\pi_k$ (constant pentru cererea k), $\quad q_k \beta \, \tau$ (depinde de transferuri), $\quad \sum_{\ell : a_{\ell p} = 1} \mu_{k,p,\ell}$ (depinde d

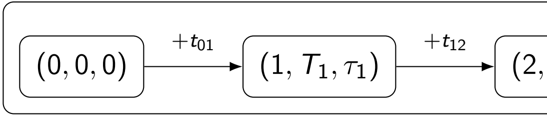
Termenii non-aditivi se tratează fie prin extinderea stării (ex. includerea liniei curente), fie prin evaluare la final.

Ilustrație: graf fizic vs. graf de stări (SPPRC clasic)

Graf fizic $G = (V, E)$



Graf de stări: noduri (v, T, τ)



Un nod din graful de stări reprezintă *aceeași stație* dar cu resurse diferite.
Fezabilitatea impune $T \leq T_k^{\max}$ și $\tau \leq \tau_k^{\max}$.

Rezolvare pricing: algoritm de etichetare (dominanță)

Idee

Menținem pentru fiecare nod v un set de etichete (labels) de forma (T, τ, C). Extindem etichetele pe arce și eliminăm cele dominate.

Dominanță (criteriu clasic)

O etichetă $L_1 = (T_1, \tau_1, C_1)$ domină $L_2 = (T_2, \tau_2, C_2)$ dacă:

$$T_1 \leq T_2, \quad \tau_1 \leq \tau_2, \quad C_1 \leq C_2,$$

cu cel puțin o inegalitate strictă.

- Doar etichetele nedominated sunt păstrate (pruning eficient).
- Dacă se găsește o rută p cu $\bar{c}_{k,p} < 0$, ea devine coloană candidat în RMP.

Algoritmul de *Column Generation*

Algorithm 1 Column Generation (schemă standard)

Inițializează un set fezabil de rute $\mathcal{P}_k^0$ pentru fiecare k (ex.: shortest-path după timp)
repeat
    Rezolvă RMP pe $(\mathcal{L}^t, \mathcal{P}_k^t)$ și extrage dualele (în special π și μ)
    Pentru fiecare $k \in \mathcal{K}$, rezolvă pricing (SPPRC) și obține o rută $p^\star$
    if există k cu $\bar{c}_{k,p^\star} < 0$ **then**
        Adaugă în RMP coloanele găsite și actualizează $\mathcal{P}_k^{t+1}$
    end if
until nu mai există coloane cu cost redus negativ

Euristici standard de accelerare

Adăugare multiplă

Se adaugă toate (sau primele M) coloane cu $\bar{c} < 0$, reducând numărul de rezolvări RMP.

Restrângere controlată în pricing

Se pot folosi praguri mai stricte:

$$T_{k,p} \leq \hat{T}_k < T_k^{\max}, \qquad \tau_{k,p} \leq \hat{\tau}_k < \tau_k^{\max}.$$

Stabilizare duală

Tehnici de stabilizare reduc oscilațiile în duale (îmbunătățesc convergența practică).

Justificare: echivalența cu simplex

Fapt cheie

Column generation este metoda simplex aplicată LP-ului complet, dar cu generare întârziată a variabilelor $y_{k,p}$.

Condiția de optimalitate

Dacă pentru fiecare cerere k pricing nu produce coloane cu $\bar{c}_{k,p} < 0$, atunci soluția curentă a RMP este optimă pentru LP-ul complet.

Relaxare LP, fracționaritate și *branch-and-price*

Modelul este o relaxare LP: x_ℓ continuu, $y_{k,p} \in [0,1] \Rightarrow$ alocare fracționară.
În general, soluția optimă LP nu este integrală.

Soluție integrală

Se utilizează *branch-and-price*: ramificare + column generation în fiecare nod al arborelui.

Concluzii

Variabilele pe rute duc la dimensiune foarte mare $\Rightarrow$ CG este natural.
Dualele din RMP definesc costul redus; pricing caută coloane cu $\bar{c} < 0$.
Pricing-ul este SPPRC: *shortest path* pe graf de stări (v, T, τ).
Pentru integritate: *branch-and-price*.