

Optimalizace rozložení skladu

EA se surrogate modely, feature selection a ensemble metody

David Zeman, Petr Šmíd, Dita Chabičovská

MFF UK

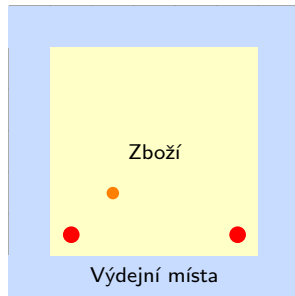
2026

Problém

- Sklad 9×9 , 49 pozic zboží, 4 agenti
- Distribuce objednávek: Zipf 1.2
- Cíl: najít layout maximalizující throughput
- Prostor: $49! \approx 6 \times 10^{62}$ permutací
- Jedna simulace: 2–3s

Surrogate model:

- 40 features z layoutu \rightarrow predikce fitness
- Hybridní EA: surrogate + občasné reálné evaluace



Červená = populární zboží

MAPD – Multi-Agent Pickup and Delivery

Problém:

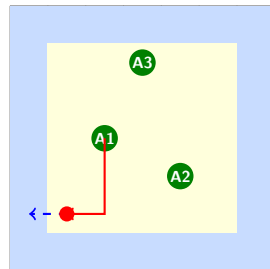
- k agentů, nekonečný proud úkolů
- Úkol = vyzvednutí zboží + doručení na výdej
- Agenti se nesmí srazit
- Cíl: maximalizovat throughput

Náš přístup:

- **Token Passing** – decentralizované přidělování úkolů
- **A* planner** – hledání cesty pro každého agenta
- Rezervační tabulka – prevence kolizí

Fitness funkce:

- $\text{fitness} = \frac{\text{dokončené úkoly}}{\text{čas simulace}}$



Zelená = agenti
Červená = cíl (pickup)
Modrá = cesta k výdeji

Extrahované features (40)

Distance-based (10):

- Vážený průměr/std vzdálenosti k okraji
- Min/max/avg vzdálenost top položek
- Vážená kvadratická vzdálenost

Position-based (8):

- Počet populárních na okraji
- Podíl v rozích
- Kvadrantové rozložení (NW, NE, SW, SE)

Distribution-based (6):

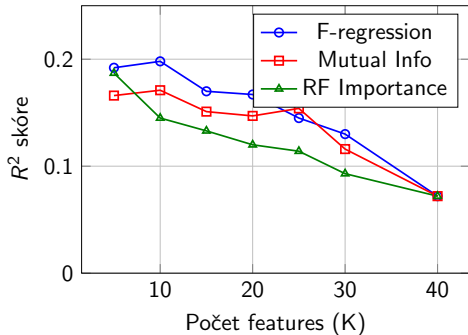
- Entropie, Gini koeficient
- Prob. hmota na okraji/středu
- Medián, 90. percentil, skewness

Congestion/clustering (10):

- Skóre kongesce, hotspot
- Lokální hustota, dispersion
- Bottleneck, path overlap
- Spread X/Y, quadrant balance

Delivery-based (6):

- Vzdálenost k výdejním místům
- Počet dostupných delivery points



Top 5 features (RF importance):

- 1 edge_prob_mass (0.142)
- 2 dist_gini (0.059)
- 3 quad_SE (0.050)
- 4 spread_x (0.050)
- 5 quad_NW (0.044)

Výsledek

- Optimum: **K = 10** features
- R^2 : 0.072 \rightarrow **0.198** (+175%)

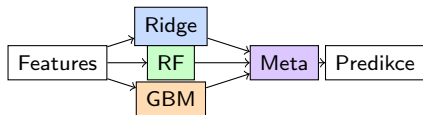
Ensemble Surrogate

Voting Ensemble:

- Ridge + RF + GBM
- Výstup = průměr predikcí

Stacking Ensemble:

- Ridge + RF + GBM
- Meta-learner kombinuje predikce



Porovnání R^2 (CV):

Model	R^2
Voting (10 feat.)	0.166
GP (40 feat.)	0.152
Ridge (10 feat.)	0.145
RF (40 feat.)	0.145
Stacking (10 feat.)	0.123
Ridge (40 feat.)	0.072

Metoda	Fitness	Čas
Random (best)	0.48	–
Greedy	0.44	–
EA bez surrogate	0.55	43 min
EA + Voting	0.56	7.5 min
EA + Stacking	0.58	10.5 min

Reálných evaluací: 1020 → **168** (–83.5%)

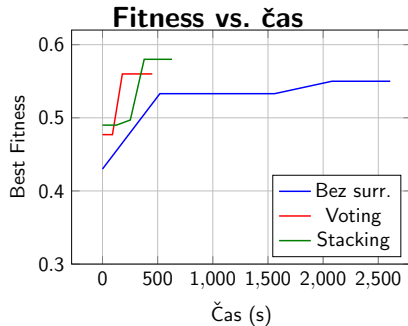
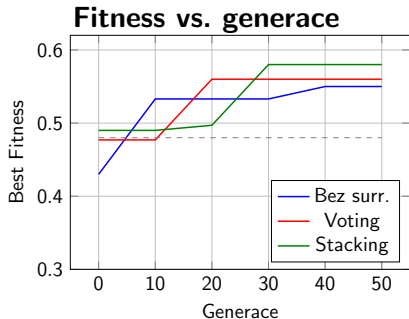
Nejlepší: EA + Stacking

- Fitness **0.58** (+5.5% vs EA)
- **4× rychlejší** než EA bez surrogate
- +21% oproti Greedy baseline

Klíčový závěr

Ensemble surrogate + feature selection = lepší výsledky za méně času

Konvergence



Surrogate model výrazně zrychluje EA optimalizaci

Co jsme udělali

- Feature selection: 40 → 10 features
- Ensemble surrogate modely
- Integrace do EA optimalizace

Co jsme dosáhli

- **4–6× rychlejší** optimalizace
- **83% méně** simulací
- **Lepší fitness** než baseline

Doporučení

Stacking ensemble – nejlepší kvalita řešení (fitness 0.58)

Voting ensemble – nejrychlejší varianta (5.8× speedup)