Information Relaxation Bounds for Dynamic Decisions on Time-Space Networks

Lavanya Marla, Mozhde Bagheri Hosseini
University of Illinois at Urbana-Champaign, lavanyam,mozhdeb2@illinois.edu

Motivation. The near real-time dynamic management of large service networks in logistics, transportation, energy, on-demand service delivery, communications and defense applications, has significant importance due to the high degree of uncertainty observed in these systems. The importance of dynamically managing systems in applications involving airline recovery (Barnhart and Vaze (2015), vehicle routing (Secomandi 2001), pickup and delivery (Psaraftis, Wen, and Kontovas 2016), job shop scheduling (Meloni, Pacciarelli, and Pranzo 2004), and energy (Lorca et al. 2016) is well-documented, while also acknowledging that these problems are NP-hard due to uncertainty (Kouki, Chaar, and Ksouri 2009).

In these settings, information about the system is revealed over time and resource allocations need to be changed to match the information corresponding to new demands or modified network attributes such as disruptions. Methods to improve the performance of these systems often consist of the design of practical heuristics to dynamically manage these networks. Such heuristics often do not have performance guarantees, both because uncertainty is only characterized partially and/or probabilistically, and because the problem is often solved approximately to ensure tractability.

We focus in this work on constructing efficient performance evaluation methods for large-scale spatio-temporal dynamic optimization problems, typically modeled on time-space networks (Ford and Fulkerson 1962, Skutella 2009). In the past, the performance evaluation of such dynamic solution methods has received limited attention. Often these methods are compared against regret-based bounds or omniscient information-based bounds, such as in online optimization and machine learning literature (Van Hentenryck and Bent 2009, Flajolet and Jaillet 2017).

Approach. Our work builds on Brown, Smith, and Sun (2010), who present an approach for information-based bounds in dynamic programs. Their approach relaxes non-anticipativity constraints that restrict the information available to dynamic programs, and instead, penalize the additional information available to the decision-maker. This is a dual-based approach for which the authors develop weak duality, strong duality and complementary slackness results. They also discuss that this approach is effective in cases where the penalty function is easy to formulate and efficient to compute a priori.

In the context of resource allocation on time-space networks, we advance Brown, Smith, and Sun (2010)'s method as follows. We define penalty functions for the case of non-anticipativity constraints as a difference between two generating functions, both defined on novel time-space network

representations. The new time-space network representation simultaneously captures *both* non-anticipative information settings (information revealed over time), and anticipative information (future information known in advance). The first generating function, defined on a time time-space network, captures the non-anticipative information relaxation case; and the second generating function is defined on a time-space network that captures the non-anticipative case. These time-space networks are defined exactly at each decision-making point in the primal dynamic programming or heuristic approach, and the difference between the solutions on these time-space networks provides the penalty function. However, instead of a priori computing the penalty function at all possible system states, we provide an algorithmic framework that allows the generating function, and the associated penalty, to be computed as part of the optimization model. We thus avoid the necessity of enumerating all possible system states of the system to compute performance bounds, and formulate the bound as a single, albeit large, integer program.

We discuss our approach primarily in the context of airline disruption management; however, our approach is easily extended to the broad class of problems involving dynamic multi-commodity flows. Figure 1 shows the airline recovery process, modeled as a rolling-horizon framework.

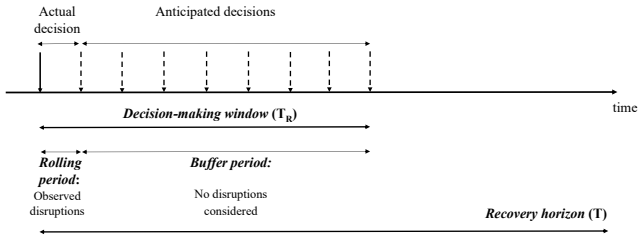


Figure 1 Rolling horizon framework for dynamic disruption management.

Let $J_t(\boldsymbol{R^t})$ be the optimal recovery-cost-to-go in period t, with the standard dynamic program with only current information (non-anticipative filtration $\mathbb{F}$) is:

$$J^t_{\mathbb{F}}(\boldsymbol{U^0},...,\boldsymbol{U^{t-1}}) = \inf_{\boldsymbol{U^t}\in\boldsymbol{\mathcal{U}}_{\mathbb{F}}(\boldsymbol{U^0},...,\boldsymbol{U^{t-1}})} \{\boldsymbol{C^t}(\boldsymbol{U^0},...,\boldsymbol{U^t}) + \mathbb{E}[J^{t+1}_{\mathbb{F}}(\boldsymbol{U^0},...,\boldsymbol{U^{t-1}})|\mathcal{F}^t]\} \tag{1}$$

A decision-maker with access to some exact future information (anticipative filtration $\mathbb{G}$) is:

$$J^t_{\mathbb{G}}(\boldsymbol{U^0},...,\boldsymbol{U^{t-1}}) = \inf_{\boldsymbol{U^t}\in\boldsymbol{\mathcal{U}}_{\mathbb{G}}(\boldsymbol{U^0},...,\boldsymbol{U^{t-1}})} \{\boldsymbol{C^t}(\boldsymbol{U^0},...,\boldsymbol{U^t}) + \mathbb{E}[J^{t+1}_{\mathbb{G}}(\boldsymbol{U^0},...,\boldsymbol{U^{t-1}})|\mathcal{G}^t]\}, \tag{2}$$

and provides a lower bound to the cost achievable by (1). Define a dual-feasible penalty function P_t at time t as one that satisfies, for all *non-anticipative actions* $\boldsymbol{U}$ under $\mathbb{F}$, satisfies $\mathbb{E}[P^t(\boldsymbol{U})|\mathbb{F}^t] \leq 0 \forall t$. Then, a 'weak duality' statement can be stated as follows.

$$\mathbb{E}[\boldsymbol{C}(\boldsymbol{U}_{\mathbb{F}})] \geq \mathbb{E}[\boldsymbol{C}(\boldsymbol{U}_{\mathbb{F}}) + \boldsymbol{P}(\boldsymbol{U_F})] \geq \inf_{\boldsymbol{U}_{\mathbb{G}}\in\boldsymbol{\mathcal{U}}_{\mathbb{G}}} \mathbb{E}[C(\boldsymbol{U}_{\mathbb{G}})) + P(\boldsymbol{U}_{\mathbb{G}})] \tag{3}$$

Additionally, define $P^t(\boldsymbol{U}) = P_1^t(\boldsymbol{U}) - P_2^t(\boldsymbol{U})$, where $P_1^t(\boldsymbol{U}) = \min_{\hat{\boldsymbol{U}} \in \boldsymbol{\mathcal{U}}_{\mathbb{F}}} \mathbb{E}[s^t(\hat{\boldsymbol{U}})|\mathcal{F}^t] = \mathbb{E}[p_1^t]$ and $P_2^t(\boldsymbol{U}) = \min_{\bar{\boldsymbol{U}} \in \boldsymbol{\mathcal{U}}_{\mathbb{G}}} \mathbb{E}[s^t(\bar{\boldsymbol{U}})|\mathcal{G}^t] = \mathbb{E}[p_2^t]$; such that $p_1^t = \min_{\hat{\boldsymbol{U}} \in \boldsymbol{\mathcal{U}}_{\mathbb{F}}} s^t(\hat{\boldsymbol{U}})|\mathcal{F}^t$, $p_2^t = \min_{\hat{\boldsymbol{U}} \in \boldsymbol{\mathcal{U}}_{\mathbb{F}}} s^t(\hat{\boldsymbol{U}})|\mathcal{G}^t$ and $P(\boldsymbol{U}) = \sum_{t=0}^{T} P^t(\boldsymbol{U})$.

This penalty function satisfies: (i) For all non-anticipative actions $\boldsymbol{U'}$ under $\mathbb{F}$, we have $\mathbb{E}[P_t(\boldsymbol{U})|\mathbb{F}_t] = 0 \,\forall\, t$, and (ii)$(P_0(\boldsymbol{U}), ..., P_T(\boldsymbol{U}))$ is adapted to $\mathbb{G}$ and P_t depends only on the first $t+1$ actions $(\boldsymbol{U}^0, ..., \boldsymbol{U}^t)$ of $\boldsymbol{U}$. Brown, Smith, and Sun (2010) continue this analysis to finally describe the bound by the following dynamic program:

$$J_{\mathbb{G}}^t(\boldsymbol{U}_0, ..., \boldsymbol{U}_{t-1})$$

$$= \inf_{\boldsymbol{U^t} \in \boldsymbol{\mathcal{U}^t}(\boldsymbol{U^0}, ..., \boldsymbol{U^{t-1}})} \left\{ \boldsymbol{C^t}(\boldsymbol{U^0}, ..., \boldsymbol{U^t}) + P_1^t(\boldsymbol{U^0}, ..., \boldsymbol{U^t}) - P_2^t(\boldsymbol{U^0}, ..., \boldsymbol{U^t}) + \mathbb{E}[J^{t+1}(\boldsymbol{U^0}, ..., \boldsymbol{U^{t-1}})|\mathcal{G}^t] \right\} \tag{4}$$

$$= \inf_{\boldsymbol{U^t} \in \boldsymbol{\mathcal{U}^t}(\boldsymbol{U^0}, ..., \boldsymbol{U^{t-1}})} \Big\{ \boldsymbol{C^t}(\boldsymbol{U^0}, ..., \boldsymbol{U^t}) + \min_{\hat{\boldsymbol{U}} \in \boldsymbol{\mathcal{U}}_{\mathbb{F}}} \mathbb{E}[\boldsymbol{C}(\hat{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \hat{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{F}] - \min_{\bar{\boldsymbol{U}} \in \boldsymbol{\mathcal{U}}_{\mathbb{G}}} \mathbb{E}[\boldsymbol{C}(\bar{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \bar{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{G}]$$
$$+ \mathbb{E}[J^{t+1}(\boldsymbol{U^0}, ..., \boldsymbol{U^{t-1}})|\mathcal{G}^t] \Big\} \tag{5}$$

In this work, we modify the above bound in (5) to the following:

$$J_{\mathbb{G}}^t(\boldsymbol{U}^0, ..., \boldsymbol{U}^{t-1})$$

$$\geq \inf_{\boldsymbol{U^t} \in \boldsymbol{\mathcal{U}^t}(\boldsymbol{U^0}, ..., \boldsymbol{U^{t-1}})} \Big\{ \boldsymbol{C^t}(\boldsymbol{U^0}, ..., \boldsymbol{U^t}) + \min_{\hat{\boldsymbol{U}} \in \boldsymbol{\mathcal{U}}_{\mathbb{F}}} \mathbb{E}[\boldsymbol{C}(\hat{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \hat{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{F}] - \mathbb{E}_{\bar{\boldsymbol{U}} \in \boldsymbol{\mathcal{U}}_{\mathbb{G}}}[\boldsymbol{C}(\bar{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \bar{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{G}]$$
$$+ \mathbb{E}[J^{t+1}(\boldsymbol{U^0}, ..., \boldsymbol{U^{t-1}})|\mathcal{G}^t] \Big\}, \tag{6}$$

$$= \inf_{\boldsymbol{U^t} \in \boldsymbol{\mathcal{U}^t}(\boldsymbol{U^0}, ..., \boldsymbol{U^{t-1}})} \Big\{ \boldsymbol{C^t}(\boldsymbol{U^0}, ..., \boldsymbol{U^t}) + \mathbb{E}_{\hat{\boldsymbol{U}} \in \boldsymbol{\mathcal{U}}_{\mathbb{F}}}[\boldsymbol{C}(\hat{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \hat{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{F}] - \mathbb{E}_{\bar{\boldsymbol{U}} \in \boldsymbol{\mathcal{U}}_{\mathbb{G}}}[\boldsymbol{C}(\bar{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \bar{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{G}]$$
$$+ \mathbb{E}[J^{t+1}(\boldsymbol{U^0}, ..., \boldsymbol{U^{t-1}})|\mathcal{G}^t] \Big\}, \tag{7}$$

$$\text{where } \mathbb{E}[\boldsymbol{C}(\hat{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \hat{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{F}] \geq \mathbb{E}[\boldsymbol{C}(\bar{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \bar{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{G}], \text{ making it a feasible penalty.} \tag{8}$$

The third term in (6) arises from replacing p_2^t by a *good feasible solution* $\bar{\boldsymbol{U}}$ under the anticipative filtration $\mathbb{G}$ instead of the optimal anticipative solution in (5). (6) will still be a lower bound, but to avoid the bound being worse than the anticipative bound, we will ensure that $\min_{\hat{\boldsymbol{U}}} \mathbb{E}[\boldsymbol{C}(\hat{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \hat{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{F}] \geq \mathbb{E}[\boldsymbol{C}(\bar{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \bar{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{G}]$, so that we have a meaningful lower bound.

We now describe the functions $\min_{\hat{\boldsymbol{U}}} \mathbb{E}[\boldsymbol{C}(\hat{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \hat{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{F}]$ and $\mathbb{E}[\boldsymbol{C}(\bar{\boldsymbol{U}}^{\boldsymbol{t+1}}, ..., \bar{\boldsymbol{U}}^{\boldsymbol{T}})|\mathbb{G}]$ as optimization problems on specially constructed time-space networks that capture both non-anticipative and anticipative decisions simultaneously, and capture the decisions through integer programs on these time-space networks. We omit descriptions for brevity, but provide them in the full paper.

Results. We evaluate our bound against (i) the omniscient lower bound that a priori knows all future disruption information, and (ii) an enhanced recovery approach called SRPDM from Lee, Marla, and Jacquillat (2019) that adopts a look-ahead heuristic, and the airline's baseline.

We consider 3 fleet types: A319, A320 and B752 of a major US airline's network. Table 1 shows that as expected, the omniscient lower bound results in a (loose) lower bound to the baseline airline operations that do not account for any disruption information beyond the current rolling period. The look-ahead SRPDM model captures about the half of the average benefit from the omniscient bound with a gap of 6.07%. Our lower bound based on penalized information relaxation tightens the lower bound, and reduce this gap bridged by the SRPDM further to 3.52%. In other words, the SRPDM model bridges, on average, 31% of the gap between the baseline and our penalized information bound. Comparatively, it only bridges 18% of the gap between the baseline and the omniscient (unpenalized) lower bound. This demonstrates the value of the SRPDM heuristic in finding good proactive recovery solutions, and the penalized information relaxation bound in evaluating such heuristics.

Table 1 Experimental results for different bounds for the A319 network

Model	Total cost ($)			% improvement over baseline		
	(avg.)	(min.)	(max.)	(avg.)	(min.)	(max.)
Industry Baseline	1,225,952	1,133,878	1,358,403	–	–	–
SRPDM (Lee, Marla, and Jacquillat 2019)	1,195,413	1,121,481	1,293,446	2.49%	1.09%	4.78%
Lower Bound B^{p-omni}	1,127,439	1,095,887	1148650	8.03%	3.52%	15.44%
Omniscient	1,056,291	1,061,005	997,530	13.8%	6.07%	26.56%

Conclusions. Our work introduces the notion of performance bounds for dynamic decisions on time-space networks, and shows that while primal heuristics based on rolling horizon algorithms, look-ahead predictions and other heuristics are often computed, dual bounds that upper bound the performance of any look-ahead algorithms can also be efficiently constructed. Moreover, in the age of data-driven and machine-learning-based predicstions, our method also helps compute the value of improved predictions – specifically, if a heuristic approach performs 'close enough' to the performance bound, we can then bound the value of forecasting, helping decision-makers decide if there is value to better predictions incorporated in their optimization models.

References

Barnhart C, Vaze V, 2015 *Irregular operations: Schedule recovery and robustness. The Global Airline Industry* 253–274.

Brown DB, Smith J, Sun P, 2010 *Information relaxations and duality in stochastic dynamic programs. Operations Research* 58(4-Part-1):785–801.

Flajolet A, Jaillet P, 2017 *Logarithmic regret bounds for bandits with knapsacks. Working Paper* .

Ford LR, Fulkerson DR, 1962 *Flows in Networks* (Princeton University Press).

Kouki Z, Chaar BF, Ksouri M, 2009 *Extended cnp framework for the dynamic pickup and delivery problem solving.* Iliadis, Maglogiann, Tsoumakasis, Vlahavas, Bramer, eds., *Artificial Intelligence Applications and Innovations III*, 61–71 (Springer US).

Lee J, Marla L, Jacquillat A, 2019 *Dynamic airline disruption management under airport operating uncertainty. Transportation Science, Minor Revision* .

Lorca A, Sun XA, Litvinov E, Zheng T, 2016 *Multistage adaptive robust optimization for the unit commitment problem. Operations Research* 64(1):32–51.

Meloni C, Pacciarelli D, Pranzo M, 2004 *A Rollout Metaheuristic for Job Shop Scheduling Problems. Annals of Operations Research* 131:215–235.

Psaraftis HN, Wen M, Kontovas CA, 2016 *Dynamic vehicle routing problems: Three decades and counting. Networks* 67:3–31.

Secomandi N, 2001 *A Rollout Policy for the Vehicle Routing Problem with Stochastic Demands. Operations Research* 49(5):796–802.

Skutella M, 2009 *An introduction to network flows over time.* Cook WJ, Lovász L, Vygen J, eds., *Research Trends in Combinatorial Optimization*, 451–482 (Springer).

Van Hentenryck P, Bent R, 2009 *Online Stochastic Combinatorial Optimization* (MIT Press, Cambridge, MA).