



CAN UNCLASSIFIED



DRDC | RDDC
technologysciencetechnologie

Optimizing course scheduling with heterogenous resource constraints uncertain student demand

Cheryl Eisler

Min Jing Liu and Peter Dobias

WinterSim 2018

Date of Publication from Ext Publisher: December 2018

Defence Research and Development Canada

External Literature (P)

DRDC-RDDC-2018-P061

April 2018

CAN UNCLASSIFIED

CAN UNCLASSIFIED

IMPORTANT INFORMATIVE STATEMENTS

This document was reviewed for Controlled Goods by Defence Research and Development Canada (DRDC) using the Schedule to the *Defence Production Act*.

Disclaimer: This document is not published by the Editorial Office of Defence Research and Development Canada, an agency of the Department of National Defence of Canada but is to be catalogued in the Canadian Defence Information System (CANDIS), the national repository for Defence S&T documents. Her Majesty the Queen in Right of Canada (Department of National Defence) makes no representations or warranties, expressed or implied, of any kind whatsoever, and assumes no liability for the accuracy, reliability, completeness, currency or usefulness of any information, product, process or material included in this document. Nothing in this document should be interpreted as an endorsement for the specific use of any tool, technique or process examined in it. Any reliance on, or use of, any information, product, process or material included in this document is at the sole risk of the person so using it or relying on it. Canada does not assume any liability in respect of any damages or losses arising out of or in connection with the use of, or reliance on, any information, product, process or material included in this document.

Endorsement statement: This publication has been peer-reviewed and published by the Editorial Office of Defence Research and Development Canada, an agency of the Department of National Defence of Canada. Inquiries can be sent to: Publications.DRDC-RDDC@drcd-rddc.gc.ca.

OPTIMIZING MILITARY COURSE SCHEDULING WITH HETEROGENOUS RESOURCE CONSTRAINTS AND UNCERTAIN STUDENT DEMAND

ABSTRACT

Planning and scheduling training courses across multiple occupations within the context of a small military is a complex task due to uncertain levels of demand and rapidly shifting requirements over time. Offering too many course sessions is resource intensive and may be unjustified or unaffordable from the resource perspective. On the other hand, if an insufficient number of sessions is offered, either some personnel will not be trained or additional costs may be incurred to initiate additional last-minute sessions. Under the assumption that the historical attendance in courses is a valid predictor of demand at least in the near future, one can derive demand estimates. The objective can then be formulated as a minimization of the number of offered sessions while meeting the demand.

Keywords: Course Loading, Scheduling Optimization, Forecasting, Sub-Problem Independence.

INTRODUCTION

The problem of class scheduling optimization has been studied in detail throughout the literature (Nakasuwan et al, 1999; Wasfy et al, 2007; Winch, 2013 to name but a few), with variations that factor in resource assignment (Badri, 1996; Gunawan, 2011), timetabling (Schaerf, 1999; Hossain et al, 2007), and various degrees of preferential heuristics (Schniederjans, 1987; Dahiya, 2015). However, the bulk of the research problems are predicated on the knowledge of the student demand. Such an assumption about the incoming training requirement may be valid in settings where deadlines are set well in advance for course session registration and the population levels are sufficiently large. In contrast, in the military training setting – especially for small militaries [with less than 120,000 total personnel and less than 4,000 major war fighting vehicles (GlobalFirepower.com, 2018)] – the problem is more complex. Because of the specialized nature of military occupations, many courses are offered on a regular basis with mandated content (similar to undergraduate level studies in university); while the attendance is generally low (similar to graduate level classes). Predicting student demand can be difficult, as there are many variables that need to be considered, such as operational requirements, deployments, major equipment procurements or maintenance issues, or changes in policy at the organizational level, as well as health, fitness, personal and family commitments, and career changes at the individual level. Within a limited time-frame, past attendance can be used to derive probabilistic distributions for the future demand, eliminating the need to factor in all individual variables that affect the attendance. Once the student demand is estimated, the number of required sessions for each course can be calculated.

This study proposed a means of modeling the supply-demand relationship as it strove to answer two questions: 1) Considering uncertain student demand, how many sessions of each specific course should be offered in order to minimize late cancellations or additions, while also maintaining a low risk of not meeting the demand? 2) How should the session schedule be organized in order to maximize the effectiveness of the resource usage? The objective for answering these two questions was to ensure that resources would be allocated in an efficient manner, meeting but not significantly exceeding demands (i.e., avoiding the need to add or cancel sessions). By extension, this would simplify resource management, including enabling more efficient longer-term acquisition and maintenance planning.

Gurvich et al (2010), Zambelli et al (2009), and Eliashberg et al (2009) propose similar variations on the problem, except only one type of resource is required to handle any given demand. In the military context, the resource considerations are significantly more complex and include unique availability restrictions on multiple distinct types of required resources (such as warships, aircraft, ground vehicles or other major platforms, qualified trainers, classrooms, virtual or mock training environments, and the supporting logistics chain). Hence, the military training problem is highly

constrained with heterogeneous resources, each with varying schedule limitations. The level of constraints limits the most common methodologies developed for civilian applications from being directly applicable to the small military training system problem.

PROBLEM DEFINITION

The analyzed problem (Eisler et al, 2015) can be described as taking a list of offered courses with the required resources and constraints, and optimizing the session schedule with respect to a predefined scoring function. This problem was separated into two parts: 1) Estimate the required session demand (the number of students vary from year to year, from session to session); and 2) Develop a session schedule and resource assignment plan based on the identified interdependencies and constraints. The first part utilized historical data for student demand with Monte Carlo Simulation Optimization (MCSO). The second part was solved as a modification to the classical Job Shop Scheduling Problem (JSSP) (Graham, 1966). The problem at hand differed from the original JSSP due to the heterogeneous resource requirements. Furthermore, the model had to accommodate the possibility of insufficient resources.

COURSE LOADING OPTIMIZATION

Session Quantity Estimation

Historical course session attendance over a five-year period was used to forecast student demand on a course-by-course basis. The distribution functions for the demand were then applied in a MCSO to deliver maximum course capacity. The historical attendance was taken to be the number of students that completed each military training course in the past five years. Ordinarily, the cancelled sessions would count as having zero students; since the objective was to reduce or eliminate cancelled sessions, it was assumed that only the non-zero student loading was valid. Demand probability distribution functions with following rules were obtained for each course:

- 1 data point (i.e., course offered only once in the last five years): No distribution, the single data point was treated as a constant and used directly.
- 2 data points: An integer uniform distribution between the two sample values was used.
- 3 or 4 data points: A triangular distribution was used due to its simplicity to compute as a rough approximation for a random variable with an unknown distribution. The minimum (*min*), maximum (*max*), and most likely value (*M*) from the probability distribution (calculated as $3 * \text{mean} - \text{min} - \text{max}$, with $M < \text{min}$ set as $M = \text{min}$ and $M > \text{max}$ set as $M = \text{max}$) were used as the parameters. Resulting student demand was rounded to the nearest integer value, which also eliminated the issue of a non-zero probability occurrence for the *min* and *max* values.
- 5 data points: A Poisson distribution was used, where the mean value of the samples is taken as the event rate (λ). Similarly, resulting student demand was rounded to the nearest integer value.

While bootstrapping techniques or other non-linear functions could have been used to represent the probability distribution functions, there was insufficient data from the sample database in order to determine better fits while maintaining realistic bounds on the sample problem size. The obtained historical distributions were then used to calculate the expected session demand using MCSO as follows. Let N be the total number of considered courses, x be the total number of scheduled sessions, y be the minimum number of sessions required to enroll all students, z be the maximum number of sessions that could be offered without cancellations due to insufficient student load, and p be the total number of students requiring courses for each course i .

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \cdot \\ x_i \\ \cdot \\ x_N \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ \cdot \\ y_i \\ \cdot \\ y_N \end{pmatrix}, \mathbf{z} = \begin{pmatrix} z_1 \\ \cdot \\ z_i \\ \cdot \\ z_N \end{pmatrix}, \mathbf{p} = \begin{pmatrix} p_1 \\ \cdot \\ p_i \\ \cdot \\ p_N \end{pmatrix} \text{ where } x_i, y_i, z_i, p_i \in \mathbb{W} = \{0, 1, 2, \dots\} \quad (1)$$

for $i = 1, 2, \dots, N$

Then \mathbf{y} (the minimum number of sessions required to enroll all students) is calculated based on dividing the total number of students for a course i by n_{\max_i} (the maximum number of students per session) and \mathbf{z} (the maximum number of sessions that could be offered without cancellations due to insufficient student load) is calculated based on dividing the total number of students for a course i by n_{\min_i} (the minimum required number of students per session). Eq. (2) does make the simplifying assumption that the initial sessions have maximum attendance, and the final session contains the remainder.

$$y_i = \begin{cases} \left\lfloor \frac{p_i}{n_{\max_i}} \right\rfloor + 1; & \text{for } (p_i \bmod n_{\max_i}) \geq n_{\min_i} \\ \left\lfloor \frac{p_i}{n_{\max_i}} \right\rfloor; & \text{for } (p_i \bmod n_{\max_i}) < n_{\min_i} \end{cases}, \text{ and } z_i = \left\lfloor \frac{p_i}{n_{\min_i}} \right\rfloor \quad (2)$$

A soft constraint or scalar penalty function (c), as a function of \mathbf{x} , \mathbf{y} and \mathbf{z} , can then be introduced.

$$c(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{i=1}^N (h_i(x_i, y_i) + k_i(x_i, z_i)) \quad (3)$$

In that penalty function, \mathbf{h} represents the number of extra sessions that must be scheduled and \mathbf{k} represents cancelled sessions for each course:

$$h_i(x_i, y_i) = \begin{cases} x_i - y_i; & \text{for } z_i \geq y_i \\ 0; & \text{for } x_i < y_i \end{cases}, \text{ and } k_i(x_i, z_i) = \begin{cases} z_i - x_i; & \text{for } z_i \geq x_i \\ 0; & \text{for } z_i < x_i \end{cases} \quad (4)$$

Since \mathbf{y} and \mathbf{z} depend on the student demand \mathbf{p} , the selection of a single set of values for \mathbf{x} for all simulated values of \mathbf{y} and \mathbf{z} becomes a penalty function minimization problem. Then, the objective function C , over t trials is:

$$C = \min(\bar{c}) = \min \left(\frac{1}{t} \sum_{\mu=1}^t (c_{\mu}(\mathbf{x}, \mathbf{y}, \mathbf{z})) \right) \quad (5)$$

$$= \min \left(\frac{1}{t} \sum_{\mu=1}^t \left(\sum_{i=1}^N (h_{\mu,i}(x_i, y_{\mu,i}) + k_{\mu,i}(x_i, z_{\mu,i})) \right) \right)$$

with $h_{\mu,i}$, $k_{\mu,i}$, $y_{\mu,i}$ and $z_{\mu,i}$ being the values of h_i , k_i , y_i , and z_i in μ^{th} ($\mu = 1, 2, \dots, t$) iteration.

Denote the fractional capacity of the i^{th} course, $f_i = p_i / (n_{\max} \times x_i)$, as the total number of students scheduled for that course divided by the product of the maximum allowable number of students per session and the total number of sessions to be held for course i . A second objective function G can then be defined as:

$$G = \max(\bar{f}) \quad (6)$$

where G is the maximum overall fractional capacity, obtained by averaging the fractional capacity of all courses. Assuming for simplicity that the demand for a course i is independent of all other courses, G and C can be decomposed into a set of sub-problems for each course as

$$\begin{aligned} G = \max(\bar{f}) &= \max\left(\frac{1}{t} \sum_{\mu=1}^t \left(\frac{1}{N} \sum_{i=1}^N (f_{\mu,i})\right)\right) = \frac{1}{tN} \times \max\left(\sum_{\mu=1}^t \left(\sum_{i=1}^N (f_{\mu,i})\right)\right) \\ &= \frac{1}{tN} \times \max\left(\sum_{\mu=1}^t (f_{\mu,1} + f_{\mu,2} + \dots + f_{\mu,N})\right) \\ &= \frac{1}{tN} \times \max\left(\sum_{\mu=1}^t (f_{\mu,1}) + \sum_{\mu=1}^t (f_{\mu,2}) + \dots + \sum_{\mu=1}^t (f_{\mu,N})\right) \\ &= \frac{1}{tN} \times \left(\max\left(\sum_{\mu=1}^t (f_{\mu,1})\right) + \max\left(\sum_{\mu=1}^t (f_{\mu,2})\right) + \dots + \max\left(\sum_{\mu=1}^t (f_{\mu,N})\right)\right) \end{aligned} \quad (7)$$

with

$$\min(c(\mathbf{x}, \mathbf{y}, \mathbf{z})) = \min\left(\frac{1}{t} \sum_{\mu=1}^t (c_{\mu}(\mathbf{x}, \mathbf{y}, \mathbf{z}))\right) = \frac{1}{t} \min(c_1) + \frac{1}{t} \min(c_2) + \dots + \frac{1}{t} \min(c_t) \quad (8)$$

The value of c_i increases when $x_i < \min(y')$ or $x_i > \max(z')$, where \mathbf{y}' is the vector of minimum and \mathbf{z}' of maximum acceptable number of sessions. Hence, only the solutions $\mathbf{x} \in [\min(\mathbf{y}'), \max(\mathbf{z}')] can minimize the penalty functions. Thus, the optimal solution that satisfies the scalar penalty function must lie between the extrema for all samples. Since \mathbf{x} must be integer, it is simpler to enumerate through the set of possible values that could theoretically satisfy all of the samples generated. The algorithm for estimating the session quantities for the i^{th} course is:$

Step 1: Generate t samples of expected student demand using the probability distribution function associated with the i^{th} course.

Step 2: For each sample (\mathbf{x}) of expected student demand, compute corresponding \mathbf{y} , \mathbf{z} , and \mathbf{p} .

Step 3: Compute $[\min(\mathbf{y}'), \max(\mathbf{z}')] for all t samples. For each integer $\mathbf{x} \in [\min(\mathbf{y}'), \max(\mathbf{z}')] calculate c'_i .$$

The final solution is then $\mathbf{c} = \{\min(c'_i), \text{ for all } t\}$.

Resource Optimization and Session Scheduling

Once a required number of sessions is determined, and given their durations and resource availability, the optimization problem can be summarized as finding the best choice of start dates with respect to some objective function within resource constraints. In order to reduce required computational time, resource selection was separated from the main scheduling problem. Given a start date proposed by an iteration of the scheduling algorithm, resources for each session are selected based on how full their overall schedules are. The loop is recursive; the schedule for each resource is updated and used as input to schedule resources for the next session.

The schedule optimization loop used the Frontline Systems' evolutionary algorithm (Frontline Systems Inc., n.d.) until predefined termination conditions were met. The primary decision variables were the session start dates; the initial value was set to 1 January for all. The start dates were subject to minimum and maximum value constraints (must fall within the modeled year), and had to be

integer. In order to define the objective function, the schedule capacity fraction cf was defined for one day of the schedule for all courses and sessions as:

$$cf = \begin{cases} \sum_{i=1}^N \sum_s cf_s - 1; & \text{for } \sum_{i=1}^N \sum_s cf_s \geq 1 \\ 1; & \text{for } \sum_{i=1}^N \sum_s cf_s = 0 \end{cases} \quad (9)$$

where the session capacity fraction, $cf_s = T_s/T_{Total}$ is the fraction of schedule slots T , that are currently assigned to a particular session. The capacity fraction cf is symmetric, with global minimum at $cf = 1$. Hence an empty schedule slot ($cf = 0$) is scored as equally undesirable as a schedule conflict. The schedule capacity fraction is then summed over an entire schedule \mathbf{A} to obtain the aggregate objective function F_{cf} .

$$F_{cf}(\mathbf{A}) = \sum_{d=1}^l \sum_{e=1}^m cf(\mathbf{A}(d, e)) \quad (10)$$

Where l and m represent time units; in this case, days per week and week number. The objective function (10) incorporated schedule conflicts (if any) and could, in the future, be expanded to include resource costs in the form of a weighted sum objective (Kim et al, 2005).

Sample Performance

A test simulation considering $N = 39$ military training courses over $l = 50$ week time span ($m = 7$ days), with 113 resources separated into 9 different pools, was conducted. The provided resource pool was very limited; this led to a non-feasible solution due to non-zero resource conflict. However, the conflicts could be resolved by relaxing the resource requirements (e.g., considering some resources as optional), or by providing additional resources. Table 1 shows a summary of the results of the generated schedule from the simulation. These results suggest a course schedule leading to significant savings in additional and cancelled sessions when compared to a past schedule (Eisler et al., 2015) could be developed.

Table 1. Case study results of the generated schedule.

Output	Value
Constraint Metric (C)	1,464
Additional Sessions Required ($\sum_i^N h_i$)	1,004
Sessions Cancelled ($\sum_i^N k_i$)	460
Aggregate Objective Function (F_{cf})	0.498
Total Sessions ($\sum_i^N x_i$)	99

CONCLUSIONS

A proposed model has been used to analyze the small military course training supply and demand relationship, and identify the optimal number of sessions that should be offered in order to meet expected student demand and minimize course cancellations and surplus. Due to the likely reasonable bound on minimum and maximum session sizes (for example, it is highly unlikely that a single session will have between 10 and 1,000 students – constraints on classroom size and/or teaching resource ratios would not permit), the enumerated approach does not become overly costly to

compute. This would extrapolate to potential annual cost savings in a large training program. In addition, tailoring the program delivery structure in a way that reduces the number of cancelled sessions, while offering sessions specifically at the times when limited assets (such as ships or major training simulators) are available to support them, would enable more efficient demand and supply forecasting, possibly leading to further savings. Due to the possibility of postings, assignments, and deployments in military careers – in addition to the usual family- or job-related responsibilities, health concerns, or other commitments that may interrupt training – the drop-out rate may be quite high, and should be a major consideration for future work.

REFERENCES

- Badri, M.A. (1996). A two-stage multiobjective scheduling model for [faculty-course-time] assignments. *European Journal of Operational Research*, 94, 16-28.
- Dahiya, S. (2015). *Course scheduling with preference optimization* (M. A. Sc. Thesis). The Pennsylvania State University Graduate School, University Park, Pennsylvania. (<https://turing.cs.hbg.psu.edu/mspapers/sources/siddharth-dahiya.pdf>, accessed 25 November 2016).
- Eisler, C., Dobias, P., and Lu., O. (2015). *Course loading optimisation*. Paper presented at the 33rd International Symposium on Military Operational Research, Barrington Digital Library at Cranfield University. (http://www.ismor.com/32ismor_archive/papers/pdf/32ismor_course-loading_paper.pdf, accessed 25 November 2016).
- Eliashberg, J, Hegie, Q., Ho, J., Huisman, D., Miller, S.J., Swami, S., Weinberg, C.B., and Weirenga, B. (2009). Demand-Driven Scheduling on Movies in a Multiplex. *International Journal of Research in Marketing*, 26, 75-88.
- Frontline Systems Inc. Excel Solver, Optimization Software, Monte Carlo Simulation, Data Mining – Frontline Systems, (<http://www.solver.com/>, accessed 25 November 2016).
- Graham, R. (1966). Bounds for Certain Multiprocessing Anomalies. *The Bell Technical Journal*, 1563-1581.
- GlobalFirepower.com (2018), 2017 Canada Military Strength, Available at: https://www.globalfirepower.com/country-military-strength-detail.asp?country_id=canada (Access Date: 6 April 2018).
- Gunawan, A., and Ng, K.M. (2011). Solving the teacher assignment problem by two metaheuristics. *International Journal of Information Management Science*, Tamkang University, 22, 73-86.
- Gurvich, I., Luedtke, J. and Tezcan T. (2010). Staffing Call Centres With Uncertain Demand Forecasts: *A Chance-Constrained Optimization Approach*. *Management Science*, 56(7), 1093-1115.
- Hossain, S., and Zibran M.F. (2007). *A multi-phase approach to the university course timetabling problem: Proceedings of the 6th Cologne Twente Workshop on Graphs and Combinatorial Optimization* (pp. 73-76). University of Twente: Enschede.
- Kim, I. Y. and de Weck, O.L. (2005). Adaptive weighted sum method for multi-objective optimization: a new method for Pareto front generation. *Structural Multidisciplinary Optimization*, 31(2), 105-116.
- Nakasuwan, J., Srithip, P. and Komolavanij, S. (1999). Class scheduling optimization. *Thammasat International Journal of Science & Technology*, Thammasat University Press, 4(2), 88-98.
- Schaerf, A. (1999). A survey of automated timetabling. *Artificial Intelligence*, 13, 87-127.
- Schniederjans, M.J. and Kim. G.C. (1987). A goal programming model to optimize departmental preference in course assignments. *Computers & Operational Research*, 14(2), 87-96.
- Wasfy, A. and Aloul F.A. (2007). Solving the university class scheduling problem using advanced ILP techniques. Paper presented at the 4th IEEE GCC Conference.
- Winch, J.K., and Yurkiewicz, J. (2013). Student Class Scheduling with Linear Programming; Proceedings for the Northeast Region Decision Sciences Institute Annual Meeting (pp. 374-386).
- Zambelli, M.S., Luna, I. and Soares, S. (2009). Long-Term Hydropower Scheduling Based on Deterministic Nonlinear Optimization and Annual Inflow Forecasting Models: Proceedings of the BUCHAREST Power Technical Conference (pp. 1-8). IEEE.

DOCUMENT CONTROL DATA

*Security markings for the title, authors, abstract and keywords must be entered when the document is sensitive

1. ORIGINATOR (Name and address of the organization preparing the document. A DRDC Centre sponsoring a contractor's report, or tasking agency, is entered in Section 8.) DRDC – Centre for Operational Research and Analysis Defence Research and Development Canada 101 Colonel By Drive Ottawa, Ontario K1A 0K2 Canada		2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.) CAN UNCLASSIFIED
		2b. CONTROLLED GOODS NON-CONTROLLED GOODS DMC A
3. TITLE (The document title and sub-title as indicated on the title page.) Optimizing course scheduling with heterogenous resource constraints uncertain student demand		
4. AUTHORS (last name, followed by initials – ranks, titles, etc., not to be used) Eisler, C.; Min, J. L.; Dobias, P.		
5. DATE OF PUBLICATION (Month and year of publication of document.) April 2018	6a. NO. OF PAGES (Total pages, including Annexes, excluding DCD, covering and verso pages.) 6	6b. NO. OF REFS (Total references cited.) 17
7. DOCUMENT CATEGORY (e.g., Scientific Report, Contract Report, Scientific Letter.) External Literature (P)		
8. SPONSORING CENTRE (The name and address of the department project office or laboratory sponsoring the research and development.) DRDC – Centre for Operational Research and Analysis Defence Research and Development Canada 101 Colonel By Drive Ottawa, Ontario K1A 0K2 Canada		
9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)	
10a. DRDC PUBLICATION NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC-RDDC-2018-P061	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)	
11a. FUTURE DISTRIBUTION WITHIN CANADA (Approval for further dissemination of the document. Security classification must also be considered.) Public release		
11b. FUTURE DISTRIBUTION OUTSIDE CANADA (Approval for further dissemination of the document. Security classification must also be considered.)		

12. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Use semi-colon as a delimiter.)

Course Loading; Scheduling Optimization; Forecasting; Sub-Problem Independence.

13. ABSTRACT/RESUME (When available in the document, the French version of the abstract must be included here.)