# Estimating the Optimum Standard Penalty of Examination Timetables

Tayo Sodeinde and George M. White
School of Information Technology and Engineering,
University of Ottawa,
Ottawa K1N 6N5, Canada
*white@site.uottawa.ca*

**Abstract**

Examination scheduling is one of the earliest applications of computer technology to an academic problem. The problem is difficult to formulate and it is usually impossible to obtain optimum solutions, or even to estimate how far from optimum a given solution is. This gave a strong impetus to academic researchers to study the problem, given the promise of a useful application and the availability of real data from real sources.

Researchers have employed a large number of techniques in order to find better solutions. Discussions of methods and summaries of progress to various dates can be found in (Carter, 1986; Carter and Laporte, 1996) and more recently in (Qu et al., 2006).

A *feasible* exam timetable is one for which no student is required to sit for more than one exam at a time. Some of these timetables are worse than others. Several measures of the "badness" of a timetable have been proposed, such as

- the total number of consecutive exams a student must write
- the total number of consecutive exams plus the total number of exams separated by exactly one free timeslot

In 1996, a seminal paper (Carter et al., 1996, 1997) used a penalty, based on some earlier work, that is the weighted sum of course pair penalties. Two exams taken by one student separated by $n$ timeslots incurs a penalty $p_n$. The number of such penalties incurred by all the students is $w_n$. The penalty of the entire timetable is then defined to be

$$\sum_{i=1}^{5} p_i w_i$$

where $p_1 = 16, p_2 = 8, p_3 = 4, p_4 = 2, p_5 = 1$, and the summation is calculated over all students involved. The penalty so obtained is then divided by the number of students involved to get a kind of *standard penalty*. The authors also referenced a depository of 13 data sets taken from real institutions that they used to test their algorithms. The benchmarks that resulted have been used ever since as a basis of comparison.

Progress since that time has been made by many researchers who have employed many different approaches with a view of lowering the standard penalty

when using new algorithms with the same data. The results are not unlike those obtained in track and field events where athletes attempt to lower the time required to cover a specified distance, say 100 metres, using the same equipment, where nothing is changed except the training techniques, diet, discipline, etc.

This raises the question as to whether past performance can be used to forecast future results. If this is true then perhaps an analytical study of past attempts to obtain lower standard penalties can be used to predict future lower standard penalties. There may even be a possibility of estimating ultimate lower bounds and even a possibility of predicting when these bounds would be obtained.

The shape of the curve that describes the experimental "current best" value points leads to the conclusion that at some time in the future, the best penalty values will reach a limit. *i.e.*

$$\lim_{t \to \infty} \frac{dP(t)}{dt} = 0$$

where $P(t)$ is the penalty obtained at time $t$. Since the exam scheduling problem is known to be NP-hard, the form of the derivative $dP/dt$ is unknown, but it is reasonable to assume that it is some function of the current bext penalty.

$$\frac{dP}{dt} = f(P)$$

Expanding this as a Maclaurin series yields

$$\frac{dP}{dt} = f(P) = a_0 + a_1 P + a_2 P^2 + a_3 P^3 + ...$$

To simplify the form of the equation we might first approximate it as $dP/dt = a_0$. Then when P attains its limiting value, we have $dP/dt = 0$ and therefore $a_0 = 0$ The next form to consider is $dP/dt = a_1 P$ which equals 0 only if $P = 0$; this is probably not the case. The next simplest form is

$$\frac{dP}{dt} = a_1 P + a_2 P^2 = P(a_1 + a_2 P)$$

This has the desired properties. When P takes its limiting value, $dP/dt = 0$. It follows that $a_1 + a_2 P_{limit} = 0$ or $P_{limit} = -a_1/a_2$ In the literature, this equation is often written in the form

$$\frac{dP}{dt} = \beta P - \delta P^2 = P(\beta - \delta P) \tag{1}$$

This is a differential solution whose solution is

$$P(t) = \frac{\beta}{\delta + [\frac{\beta}{P_0} - \delta] e^{-\beta t}} \tag{2}$$

where $P(t)$ is the penalty obtained at time $t$ and $\beta, \delta$ and $P_0$ are adjustable parameters. When $t = 0, P(t) = P(0) = P_0$. As $t \to \infty$, $P(t) \to \frac{\beta}{\delta}$. This equation is used to describe birth-death processes and race results among other applications. Here we use it to analyse some of the published examination timetable results.

For the data set *car-s-91*, the progressive "world record" was tabulated along with the year in which the work was published. The results are shown in the table below. The records correspond to the best result (if any) published during the corresponding calendar year. If the record was broken more than once during the year, the best result was taken, The earliest result is taken from the original paper (Carter et al., 1996) and the latest result (and current champion) was published in (Burke and Bykov, 2006).

| year | 1996 | 2001 | 2002 | 2003 | 2005 | 2006 |
|---|---|---|---|---|---|---|
| penalty | 7.1 | 6.2 | 5.23 | 4.54 | 4.5 | 4.42 |

The first column of data was not used in the analysis because it was the first time that the data and the penalty used was presented to the research world. The long time before the next result was published is not representative of the interval separating the next improvements. The logistic curve was fitted to the remaining five data points using the solver tool in Microsoft Excel. The goal of the solver was to minimize the squared deviations between the published results and the fitted equation while adjusting the constants $\beta, \delta$ and $P_0$. When this is done, the limiting value of the standard penalty is calculated to be 3.57.

A similar calculation for the set *car-f-92* yields the result, 3.53 . Some sets, such as *hec-s-92*, do not provide enough data to permit the calculation.

These results should be considered preliminary. There are very few data points available and the errors on the limiting values are thought to be large. For *car-s-91* there are only five valid data points and there are three adjustable parameters in the equation to be fitted. Work is continuing on finding new published results for the penalties and calculating error bounds on the limiting values obtained.

## References


Burke EK Bykov Y (2006) Solving exam timetabling problems with the flex-deluge algorithm. In Burke EK and Rudová H (ed), *Proc. of the 6th Int. Conf. on the Practice and Theory of Automated Timetabling*, pages 370–372

Carter MW (1986) A survey of practical applications of examination timetabling algorithms. *Operations Research*, 34:193–202

Carter MW G. Laporte G (1996) Recent developments in practical examination timetabling. In Burke EK and Ross P (ed) *Lecture Notes in Computer Science*, volume 1135, pages 3–21. Springer

Carter WM, Laporte G, Lee SY (1996) Examination timetabling: Algorithmic strategies and applications. *Journal of the Operational Research Society*, 47:373–383

Carter WM, Laporte G, Lee SY (1997) Corrigendum. *Journal of the Operational Research Society*, 48:225

Qu R, Burke EK, McCollum B, Merlot LTG, Lee SY (2006) A survey of search methodologies and automated approaches for examination timetabling. Technical Report NOTT-CSTR-2006-4, University of Nottingham, School of Computer Science and IT, 2006. (to be published in J. Scheduling).