19th European Symposium on Computer Aided Process Engineering – ESCAPE19
J. Jeżowski and J. Thullie (Editors)

Combinatorial Optimisation Algorithms for Strategic Biopharmaceutical Portfolio & Capacity Management

Edmund D. George, Suzanne S. Farid

The Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK, s.farid@ucl.ac.uk

Abstract

The development of a stochastic combinatorial multi-objective optimisation framework is presented which addresses the challenge of exploring large decision spaces while also capturing the multitude of decisions, trade-offs and uncertainties involved in R&D portfolio management. Detailed process economic models were linked to models of the drug development pathway to predict the technical, financial and risk implications of alternative strategies for portfolios of biopharmaceutical drugs proceeding through development. Machine learning and evolutionary computation techniques were harnessed to evolve strategies to multi-objective optimality. An industrially-relevant case study is presented that focuses on the following decisions for a portfolio of therapeutic antibodies: the portfolio composition, the scheduling of critical development and manufacturing activities, and the involvement of third parties for these activities. The impact of budgetary constraints on the optimal set of solutions is also illustrated.

Keywords: portfolio management, capacity planning, biopharmaceutical drug development, stochastic combinatorial optimisation, multi-objective

1. Introduction

Important strategic considerations for survival and success of biopharmaceutical drug developers include how best to structure portfolios, to schedule development and manufacturing activities and to acquire access to manufacturing capacity. These portfolio and capacity management decisions are further complicated by constraints on resources such as available budget and capacity as well as uncertainties that include the risk of clinical failure. Hence, the impact of making sub-optimal decisions in this environment can be severe. Frameworks that incorporate both the problems of portfolio management and manufacturing capacity planning simultaneously exist [e.g. 1, 2] and have typically used mathematical programming methods. Due to the size and complexity of this problem, a stochastic and multi-objective combinatorial optimisation approach utilising evolutionary algorithms is explored in this paper. Such approaches facilitate the capture of the interdependent activities involved in the development of drug portfolios, along with their technical, financial and risk characteristics and hence also have the capacity to give richer insight to the decision maker on the problem itself. The overall framework is a combination of a simulation-based evaluative framework based on George *et al.* [3] and a bespoke estimation of distribution algorithm (EDA) [4] to iteratively evolve a population of candidate strategies.

2. Model Formulation

The stochastic optimization process makes use of an evaluative framework (Fig. 1) coupled with evolutionary computation. The latter makes use of an estimation of distribution algorithm (EDA) (Fig. 2) that operates the optimization procedure through the machine learning of instances of decision variables that are associated with superior strategy performance using Bayesian networks. The model was designed using C++, which offers significant savings in computational time, with MS Excel as the main graphical user interface and Visual Basic for Applications (VBA) for modules controlling the flow of data in the simulation and optimization procedures. The reward and risk metrics used to to identify top performing strategies were the profitability indicators mean positive net present value (NPV) and p(NPV>0).

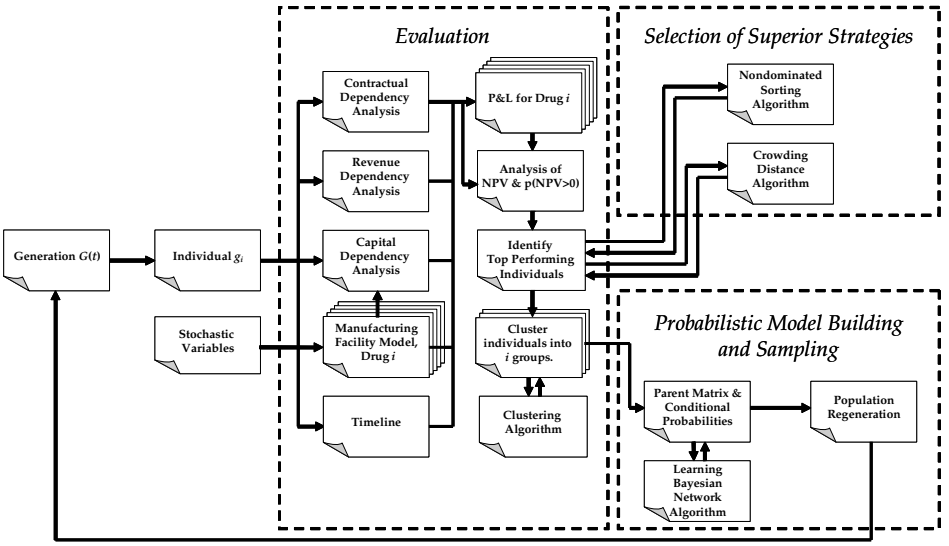

Fig. 1. Schematic of the entire simulation and optimization framework.

Estimation of Distribution Algorithm

Initial population:
Let the population of candidate solutions be $G(t)$.
$t = 1$
Randomly generate the initial population of candidate solutions, $G(1)$.
Let t_{max} be the maximum number of populations to be evolved.

For $t = 1$ to t_{max}:
Evaluation:
Simulate each g in $G(t)$ and record stochastic properties.

Selection:
Use the fast nondominated sorting and crowding distance algorithms to select the top 50% of solutions, $S(t)$, from $G(t)$.

Clustering of the objective space:
Using the k-means clustering algorithm, separate $S(t)$ into z clusters.

For each cluster, K_z:
Probabilistic model building:
Construct the Bayesian network, B_z, using a hill climbing procedure to optimize the Bayesian Dirichlet metric over B_z.

Sampling of the probabilistic model:
Generate a new set of strategies $O_z(t)$ by randomly sampling the joint probability distribution encoded by B_z. The number of strategies to be generated will be twice the original cluster size.

Population regeneration:
Generate the new population $G(t+1)$ by randomly by replacing all strategies in $G(t)$ with all strategies in each $O_z(t)$.

Fig. 2. Pseudocode for the EDA

3. Case Study Description

A hypothetical case was formulated to illustrate and examine the ability of the framework to discover optimal strategies for performance against multiple objectives in an uncertain environment. The case study concerns a biopharmaceutical company that has 10 monoclonal antibody drug candidates available for development but can only construct a drug development portfolio of limited size. It needs to know which drug candidates should be chosen, their order of their development, the timing schedule of development activities, and which corporate bodies should be assigned to each development activity. Additionally, the company is interested in knowing how solutions which are nondominated in the objective space change with the magnitude of constraint on cash flow.

The optimization model is considerate of the commercial characteristics of drug candidates, technical probabilities of success for each drug group, durations and costs

associated with various stages of the drug development and dependencies for revenue, capital expense and royalty payments. The stochastic variables included in this work are exclusively characterized by way of triangular probability distributions because of their convenience when limited sample data are available. The cash flow constraints applied to this portfolio of -$200MM, -$100MM, and -$75MM. represent the possible cash limitations of a company when funding portfolio development. If this negative cash flow is breached then clearly the company is stretched beyond the limit of finance it had intended to use.

The following settings were used for the mechanics of the optimization procedure: maximum number of iterations to be evolved (t_{max}) = 17, number of candidates in each generation at each iteration ($|G(t)|$) = 1000, number of Monte Carlo trials per candidate strategy (U) = 250, and number of superior candidate strategies in $G(t)$ at each iteration ($|S(t)|$) = 500.

4. Results and Discussion

Fig. 3 shows the final results for a five drug portfolio subject to the various constraints considered. The Pareto frontiers indicate that a negative relationship between mean positive NPV and p(NPV>0) exists for all constraints. It can also be seen in each case that for a given p(NPV>0) value the more restrictive the cash flow constraint the lower the mean positive NPV. Similarly, for a given target in mean positive NPV these constraints reduce the probability of achieving this profit.

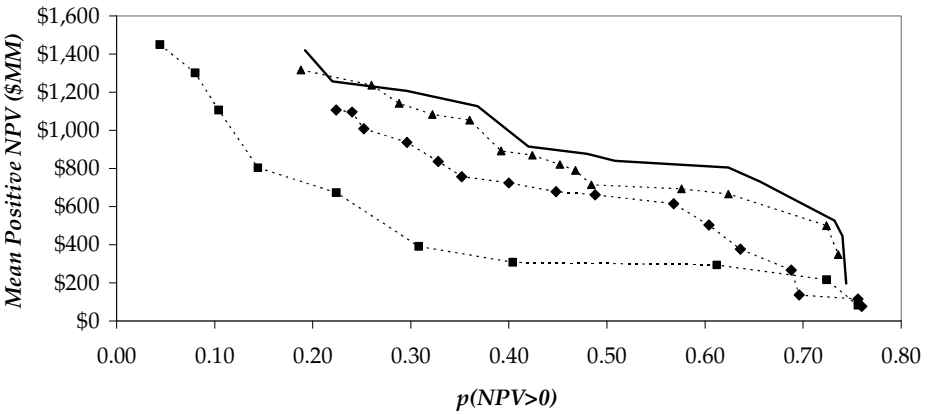

Fig. 3. Pareto frontiers of reward (mean positive NPV) versus risk (p(NPV>0)) for a five drug portfolio for the following cash flow constraints: unconstrained (-), -$200MM (▲), -$100MM(◆), -$75MM (■).

The impact of such constraints on third party, timing and drug selection strategies was analysed. An example of the most probable constituents of third party strategies for the unconstrained example alongside the -$200MM and the -$100MM constraints is given below as an example.

Third Party Strategies

When analyzing the third party strategies in the final generation for each constraint (Table 1) it is clear that the general trend in third party strategy is to develop and manufacture all drugs in-house under the presence of no constraints and then move towards the increased involvement of partners as constraints become more severe. This is understandable from the viewpoint that increasing the severity of cash flow constraints makes it imperative that more cost-effective strategies be formulated and in the problem formulation used here partners offer the most cost-effective route for drug development. For the -$200MM constraint it can be seen that in the most probable strategies partners are chosen for the clinical manufacturing stages for the first two drugs and contract manufacturers for successive drugs, indicating that this level of constraint makes in-house commercial manufacturing to be economically unattractive. The use of partners appears to be the most extensive for the -$100MM constraint. The above examples clearly show that third parties are an important resource in managing the risk and impact of failure.

Table 1. Most probable third party strategies for a five drug portfolio with different cash flow constraints for the middle cluster of strategies.

	No cash flow constraint				**- $200MM cash flow constraint**				**- $100MM cash flow constraint**			
Drug	**PI**	**PII**	**PIII**	**M**	**PI**	**PII**	**PIII**	**M**	**PI**	**PII**	**PIII**	**M**
1	*'I'*	*'I'*	*'I'*	*'I'*	*'P'*	*'P'*	*'P'*	*'P'*	*'P'*	*'P'*	*'P'*	*'P'*
2	*'I'*	*'I'*	*'I'*	*'I'*	*'I'*	*'P'*	*'P'*	*'P'*	*'I'*	*'I'*	*'C'*	*'I'*
3	*'I'*	*'I'*	*'I'*	*'I'*	*'I'*	*'C'*	*'C'*	*'P'*	*'P'*	*'P'*	*'P'*	*'P'*
4	*'I'*	*'I'*	*'I'*	*'I'*	*'C'*	*'C'*	*'C'*	*'C'*	*'P'*	*'P'*	*'P'*	*'P'*
5	*'I'*	*'I'*	*'I'*	*'I'*	*'I'*	*'C'*	*'C'*	*'C'*	*'P'*	*'P'*	*'P'*	*'P'*

Note: *'I'* – in-house activity, *'C'* – outsourced activity, *'P'* – partnered activity, PI-PIII - manufacturing for clinical phases I-III, M – manufacturing for the market phase.

5. Conclusions

A stochastic multi-objective combinatorial optimisation framework has been used to identify optimal strategies that address three key decisions simultaneously relating to portfolio and capacity management. Due to the complexity of this problem, a principle contribution of this work is in demonstrating a formulation based on techniques from evolutionary computation and machine learning employed for an efficient search of the decision space and for effective discovery of a dense and widespread Pareto frontier. It has been seen in the cases investigated here that mean positive NPV and p(NPV>0) are conflicting measures however both are desirable to the decision maker. The introduction of cash flow constraints can lead to a reduction in the expected rewards or probability of success of strategy and it directly influences the choice between in-house and external manufacturing.

Acknowledgements

The Advanced Centre for Biochemical Engineering (ACBE) and support for the Innovative Manufacturing Research Centre for Bioprocessing (IMRC) housed in the ACBE by the EPSRC are gratefully acknowledged. Financial support in the form of an

EPSRC Engineering Doctorate (EngD) studentship for E. George is also gratefully acknowledged.

References

[1] C.T. Maravelias and I.E. Grossman, Ind. Eng. Chem. Res., 40 (2001) 6147.
[2] A.A. Levis. and L.G. Papageorgiou, Comput. Chem. Eng., 28 (2004) 707.
[3] E.D. George, N.J. Titchener-Hooker, and S.S. Farid, Comput. Chem. Eng., 31 (2007) 889.
[4] M. Pelikan, D. Goldberg, and E. Cantú-Paz, Evol. Comput., 8 (2000) 311.