JOURNAL OF SPACECRAFT AND ROCKETS
Vol. 40, No. 2, March–April 2003

Multi-Objective Optimization of Communication Satellites with Two-Branch Tournament Genetic Algorithm

Rania A. Hassan* and William A. Crossley†
Purdue University, West Lafayette, Indiana 47907

In spacecraft design, many specialized state-of-the-art design tools are employed to optimize the performance of various subsystems. However, there is no structured system-level concept-definition process. Consequently, designers usually compromise some mission goals to satisfy only one of the primary design objectives. The conceptual stage of the spacecraft design process is formulated into a multi-objective discrete optimization problem. The use of multi-objective design allows the designer to evaluate different design alternatives across the whole set of design objectives. This work addresses two key design objectives for the spacecraft design process: the minimization of total launch mass and the maximization of spacecraft overall reliability. To predict values for the objective and constraint functions, a satellite design tool, which includes a satellite sizing model and a deterministic reliability model, was built and integrated with a genetic algorithm that employs a two-branch tournament to address the dual objective problem. The multi-objective approach was successful in determining sets of discrete design parameters that would minimize the launch mass as well as maximize the reliability of a geostationary communication satellite, using specified payload requirements. The designs generated by this approach appear to fall into three regions of the tradeoff space between the satellite launch mass and the satellite reliability objectives.

Nomenclature

c = penalty multiplier
f = fitness function
g = inequality constraint function
h = equality constraint function
I_{sp} = specific impulse, s
J = number of inequality constraints
K = number of equality constraints
M = number of active components
N = number of available components
N_{pop} = population size
P^* = penalty scaling factor
$\boldsymbol{x}$ = design variable vector
ΔV = change in velocity, m/s
$\boldsymbol{\phi}$ = vector of objectives
ϕ = objective function

Introduction

SYSTEMS engineering emerged as a discipline in the 1950s to manage the design and manufacture of complex systems. Systems engineering focuses on the system with a top-down approach rather than from the bottom up. Systems engineering looks at the top-level design parameters that need to be decided on early in the design process before moving to further detailed stages of the design.[1] The idea is that, if the conceptual design stage decisions are made in an appropriate way, with all of the critical trade studies performed and with all of the design objectives considered, the rest of the design process should progress smoothly with minimal costly changes either in design or manufacturing. This process, if carried out properly, should also lead to considerable savings in product cost. This paper presents a system-level multi-objective optimization approach that is applied to a commercial satellite. Using this approach, a satellite systems engineer would be able to identify tradeoffs between two competing design objectives.

Received 28 May 2002; revision received 13 September 2002; accepted for publication 5 October 2002.

*Ph.D. Candidate, School of Aeronautics and Astronautics, 1282 Grissom Hall; hassanr@ecn.purdue.edu. Student Member AIAA.

†Associate Professor, School of Aeronautics and Astronautics, 1282 Grissom Hall; crossley@ecn.purdue.edu. Senior Member AIAA.

Satellite Design

Satellite design is a complex, iterative process that involves multidisciplinary engineering expertise. An optimal spacecraft design is one that achieves four goals simultaneously: meets or exceeds mission requirements, has the highest payload to spacecraft launch mass ratio, has a high reliability, and has the lowest possible cost.

In the satellite industry, marketing specialists prepare mission requirements (payload performance) based on average market standards and pass these to engineering to start the design process. Engineers develop a conceptual design for each subsystem based largely on their experience. The integration and final selection of an overall design concept is then decided on in an iterative way by the multidisciplinary team. Such preliminary design activities, which set the goals of the whole design process, involve at best a limited evaluation of the system performance as a whole and cannot account for all possible combinations of subsystem designs. Consequently, some of the mission requirements are either jeopardized or achieved with much higher cost than planned when problems are discovered later in the design process.

Satellite design involves many interrelated design parameters; therefore, the designer has to understand the effect of each decision on the whole system. A designer can improve the design of one subsystem and simultaneously affect another subsystem negatively, which may lead to an overall enhancement or deterioration on the system level. For example, electric propulsion technology provides significant reductions in propellant mass due to its high specific impulse. However, electric propulsion requires high-power levels, which in turn increases the size and mass of the power subsystem components. Therefore, electric propulsion can introduce total system advantage only if combined with efficient lightweight power systems.[2]

Reliability as a Design Objective

In many terrestrial engineering systems, a failure can be tolerated as long as it does not jeopardize the survival of the system or safety of humans. General maintenance procedures preclude designs with complete built-in reliability. Unfortunately, this luxury cannot be afforded in space systems where on-orbit maintenance is nearly impossible; therefore, reliability is crucial to system survival.[3] In spacecraft design, mandatory high-reliability requirements tend to lead designers and decision makers toward older, conservative, oversized designs that usually entail higher cost due to multiple redundancies. Systems engineers typically apply large safety margins, believing that it will lead to high confidence levels in their designs. One of the reasons for this approach is that designers often distrust published failure data.[4]

Measures to address reliability include either fault avoidance, or fault tolerance, or both. Fault avoidance refers to oversizing parts by using high design margins to avoid any sort of failure; this technique can be used only if failure modes are well known. Fault tolerance allows for the continuation of the mission in spite of failures by employing redundancy. Fault avoidance is usually less costly than fault tolerance. Reducing cost does not necessarily mean reducing reliability. Simple missions with fewer components and subsystems are, in general, more reliable than their traditional larger spacecraft by virtue of their lower parts count.[5] The difference between what the designer wants to implement in the design in terms of reliability goals and what the design budgets can actually afford must be realized early in the design process.

For bandwidth providers in the commercial satellite industry, reliability of the spacecraft translates into availability, which, in turn, translates into revenue generation; therefore, reliability should be traded off against impact on revenue. Consequently, an accurate realistic representation of reliability is greatly needed. Same design redundancy is always applied in commercial satellite designs. The cost of same design redundancy can be significantly reduced if M-out-of-N replacement is used, in which the spare units $(N - M)$ can replace any one of the active M units. This strategy is commonly employed in the repeaters' high-power amplifiers (HPAs), battery cells, and solar array strings.

Automation of Spacecraft Design

The economics of the spacecraft industry is a major player when it comes to making design decisions. Meeting high-performance goals under constrained budgets can only be achieved through a well-structured optimized design process. There have been few efforts that introduced the concept of optimization or automation for spacecraft design, particularly in the earliest phases of design. George et al. have been working on a design tool that connects all spacecraft performance analysis tools on a network under control of a master design program.[6] Lamarra and Dunphy further developed the work described in Ref. 6 by incorporating web-based real-time collaborative interactivity.[7,8] Fukunaga et al. also used adaptive evolutionary algorithms as optimization techniques for spacecraft design.[9] The main problem in the approach adopted by these previous researchers is that, because detailed performance analysis tools are used, the project requires computing platforms of immense power. A better approach for early phases of satellite design is to build a simple, fast design tool that addresses most, if not all, of the design process aspects that are usually performed manually by systems engineers. For this tool to be effective, it has to be easily implemented using any cheap computing platform.

In 1997, Mosher[10] introduced the idea of optimization of spacecraft design in the conceptual design stage using genetic algorithms. The model included simple design estimating relationships (DERs) and cost estimating relationships. The model was limited in scope and involved only cost as a single objective function with some constraints on launch vehicle reliability. This effort was a preliminary demonstration of the applicability of evolutionary optimization methods in spacecraft conceptual design. Pullen and Parkinson have also demonstrated the successful use of evolutionary algorithms in the reliability optimization of the Gravity Probe-B spacecraft bus.

Previous work by the authors of this paper demonstrated the capability of a genetic algorithm (GA) to produce nonintuitive optimal commercial satellite system-level designs. The results described in Ref. 11 illustrate the importance of integrating payload design with overall system design in the conceptual design stage.

Multi-Objective Optimization

In a multi-objective design optimization problem, there is seldom one optimum solution as in single-objective optimization. There are usually many optimum design points, which are called the Pareto optimal set of solutions after engineer/economist Vilfredo Pareto.[12] For each design solution in this set, there is no other solution that can be found that is better on all objectives than the Pareto optimal design; in other words, these designs are nondominated.[13] If there is any solution that is better on one or two objectives than a Pareto optimal solution, then this dominated design will be worse on at least one of the remaining objectives. The Pareto front is a curve or surface that illustrates design tradeoffs between the objectives. For example, a heavy, highly reliable satellite design can be a solution on the Pareto front along with a lighter, less reliable design.

The goal in multi-objective optimal design problems is to minimize a vector function whose components are individual objectives as suggested by

$$\text{minimize}\ \phi(x) = \begin{Bmatrix} \phi_1(x) \\ \phi_2(x) \\ \vdots \\ \phi_{n_{\text{obj}}}(x) \end{Bmatrix} \tag{1}$$

Mathematically, a design with a vector of objectives $\boldsymbol{u}$ dominates a design with a vector of objectives $\boldsymbol{v}$ by meeting the conditions in

$$\text{if} \quad \forall_i\ v_i \le u_i \quad \text{and} \quad \exists_i\ v_i < u_i \qquad i = 1, 2, \ldots, n \tag{2}$$

It is very useful to find the Pareto front of design solutions because it gives the designer clear understanding of the design problem tradeoffs. However, finding this set of designs is not an easy task. Traditionally, finding the Pareto optimal set has been accomplished by using single-objective optimization algorithms while treating the rest of the objectives as design constraints or by some aggregate merit function that incorporates all of the objectives. This traditional approach requires many runs of the algorithm to find a good approximation of the Pareto front, with no guarantee that all of the designs are nondominated designs. Examples of these more traditional methods using single objective algorithms include weighted sum,[14] goal attainment,[14] ε constraint,[14] and normal-boundary intersection.[15]

Scope and Methods of Approach

This study focuses on investigating the applicability of a two-branch tournament GA for multi-objective optimization of the conceptual design of a communication satellite. Two design objectives are considered: minimizing the total satellite launch mass and minimizing the overall system risk, which is equivalent to maximizing overall system reliability. To that end, a satellite sizing code was specifically developed for use in the design optimization tool. It includes DERs that predict the performance of the system based on subsystem characteristics. A complete description of the satellite sizing tool is provided in Ref. 11. In addition, a deterministic reliability model was developed and integrated with the sizing tool to estimate the total system reliability based on subsystems options and amount of redundancies.

Many of the estimating relationships in the sizing tool are based on four recent communication satellites. The results of the suggested optimization methodology are compared to the design parameters of one of those missions. The satellite design mission that is under investigation for this work provides telecommunication services in C- and Ku-frequency bands. The existing baseline satellite for this mission employs traveling wave tube amplifiers (TWTAs) for the Ku-band repeater and TWTAs as well as solid state power amplifiers (SSPAs) for the C-band repeaters. The platform has nickel hydrogen batteries and deployable silicon solar arrays, and it employs north/south (N/S) thermal coupling. The satellite design also employs bipropellant thrusting for transfer orbit, N/S station keeping, east/west (E/W) station keeping, and attitude control.

Satellite Reliability Model

The deterministic reliability model calculates the reliability of each subsystem based on the type of technology it utilizes and the level of redundancy implemented in the design of that subsystem. The reliability model first calculates payload reliability, bus subsystems reliabilities, and then overall system reliability. Launch vehicle reliability is also incorporated.

Some of the reliability values used in this model were assumed because failure rates for some components are not available in published data. The assumptions were made based on the first author's experience in the satellite industry. For alternative technologies that are available for the same component or subsystem, the author assigned higher reliability values to the older, more commonly used technologies and lower reliability values for the relatively new technologies if no additional information was available. It is assumed that in an actual spacecraft design setting, these assumed failure rates would be replaced by better information.

A study at Goddard Space Flight Center showed an average of 1.7 failures per spacecraft during the first 30 days compared to an average of less than 0.2 failures per spacecraft per month during the following five months.[16] The high failure probability in the early stage of operation is because design problems are more likely to become apparent in the early phase of operation. The failure of any component or subsystem in the spacecraft during launch or orbit insertion has a much larger effect on the rest of the mission than failures in later phases of the mission. Therefore, this research uses the values of alternative designs' reliabilities at the end of the 30th day of on-orbit operation to compute payload and system reliability values.

For the payload of the telecommunication satellite mission used in this study, there are one Ku-band repeater and eight C-band repeaters for eight coverage areas. Each repeater's reliability is calculated based on the type of the HPA and the number of the active HPAs out of the number available on the satellite. The reliability of one HPA at the end of the first month of operation is calculated using exponential distribution as

$$R(t) = e^{-\lambda t} \tag{3}$$

The failure rate λ is 660 or 880 failures per billion amplifier operating hours for TWTA or SSPA, respectively.[17] This indicates that the reliability of a TWTA is higher than that of an SSPA.

Equation (4) is used to calculate total repeater reliability from individual HPA reliabilities for M active HPAs out of N available ones. To calculate the overall payload reliability, all repeater and antenna reliabilities are multiplied as shown in Eq. (5):

$$R_{\text{repeater}} = \sum_{i=M}^{N} \binom{N}{i} R^i (1-R)^{N-i} \tag{4}$$

$$R_{\text{payload}} = R_{\text{antennas}} \times R_{\text{repeaters}} \tag{5}$$

The bus reliability is a function of its six subsystems' reliabilities and the reliability of the spacecraft harness. The reliability of the structures subsystem is assumed to be 99%, whereas the harness reliability is assumed to be 99.9%. For this study, a reliability value of 99% was assumed for both the attitude determination and control subsystem (ADCS) and the telemetry, command, and ranging (TCR) subsystem. For this problem, both systems can either be stand-alone systems or they can both have full duplicate redundancies, the reliability of which can be calculated from the single-system reliability R_S as

$$R_{(1\,\text{out of}\,2)} = 2R_s - R_s^2 \tag{6}$$

The reliability of the electric power subsystem is the product of the solar array reliability and the battery reliability. For the solar array, it was assumed that silicon cells have higher reliability than gallium arsenide because gallium arsenide cells are relatively new in the industry compared to silicon cells. The number of additional cell strings used to replace any failed string represents redundancy for the solar arrays. In the satellite industry, a 5% additional solar array area is typically used for redundancy. As for battery reliability, nickel hydrogen cells are assumed to have higher reliabilities than nickel cadmium because they have a longer use history in commercial space applications.

The reliability of the thermal subsystem is a function of its complexity. If the thermal subsystem employs N/S thermal coupling, then its reliability value is less than the no-coupling option.[5] As for the propulsion subsystem, it is assumed that bipropellant thrusters have the highest reliability value, followed by hydrazine thrusters, arcjets, and finally the new plasma thrusters. The propulsion subsystem mentioned here is used for orbit insertion, station keeping (STK), and attitude control. The propulsion subsystem can have either one branch of thrusters or two fully redundant branches. Most commercial satellites are designed with two branches. The bus reliability and the spacecraft reliability can then be calculated as

$$R_{\text{bus}} = R_{\text{structure}} \times R_{\text{ADCS}} \times R_{\text{TCR}} \times R_{\text{power}} \times R_{\text{thermal}} \times R_{\text{propulsion}} \times R_{\text{harness}} \tag{7}$$

$$R_{\text{spacecraft}} = R_{\text{payload}} \times R_{\text{bus}} \tag{8}$$

Usually, failure during launch is covered by insurance, and the launch service provider may offer the satellite owner another free launch, but this does not reflect the actual losses due to failure. Investigations to establish responsibility for mission failure usually delay the insurance reimbursement. During this time, the satellite operator may lose revenue and customers until a replacement satellite is built and deployed, which can take about two years on average.[18] Choosing a highly reliable launch vehicle may be worth the high cost, when the risk associated with failure is considered. Therefore, the problem statement used in this work includes launch vehicle reliability as a design constraint.

The reliability model incorporates eight launch vehicles: Ariane 5, Ariane 4, Sea Launch, Proton, Delta, Atlas, Long March, and H2A. For each launcher, the transfer orbit ΔV is specified along with maximum launch mass and fairing diameter and height, which the sizing code uses to define the satellite bus geometry. Each launch vehicle has a figure of merit representing the launch vehicle reliability based on the number of successful launches out of overall number of launches until 1999,[19] as shown in Table 1. Launchers that have fewer than 50 successful launches were penalized an additional 25% below the 1999 published success rate; these are noted. This somewhat arbitrary strategy was used to discriminate against launchers that have a high-success rate with low number of launches. Other strategies could be implemented, if enough data were available.

Genetic Algorithms

The GA is a global search method that mimics the behavior observed in biological populations. The GA employs the principal of survival of the fittest in its search process, and it has been applied successfully to the design of many complex systems. The GA differs from conventional optimization methods in four different ways that make the GA well suited to the conceptual design of satellites. First, the GA works with a coding for the design parameters that allows for a combination of discrete and continuous parameters in one problem statement. Second, the GA is a population-based search technique, which results in multiple designs with good performance after each run of the GA, rather than only one solution. Third, the GA needs only fitness or objective function values; no derivatives are needed. This feature not only allows for discrete variables, but also allows for discontinuous objective and constraint functions. Additionally, this property means that the GA provides no information about optimality of the solution. Fourth, the GA employs probabilistic choices rather than deterministic rules to find new points with likely improvement. This probabilistic search technique means that the GA

Table 1 Reliability data for launch vehicles[19]

Launcher	Successful launches/ total launches	Success, %	Reliability index
Ariane 5	1/3	33	8[a]
Ariane 4	83/86	97	97
Sea launch	1/1	100	75[a]
Proton	229/260	88	88
Delta	78/80	98	98
Atlas	41/41	100	75[a]
Long march	46/54	85	85
H2A	5/6	83	58[a]

[a]Fewer than 50 successful launches; penalized additional 25%.

is likely to search across the entire design space; it will not become trapped in local minima (adapted from Ref. 20).

The variable coding that is used in the GA is particularly important for the spacecraft design problem because, in spacecraft design, the design variables can be discrete, such as the type of propulsion subsystem technology; integer, such as the number of redundant payload HPAs; or continuous, such as the solar array area. The GA advantage becomes apparent for handling discrete variables that cannot be handled by traditional calculus-based methods.

To perform its optimizationlike process, the GA employs three operators to propagate its population from one generation to another. The first operator is the "selection" operator that represents the principal of survival of the fittest. The second operator is the "crossover" operator, which represents mating in biological populations. The crossover operator propagates features of good surviving designs from the current population into the future population, which will have better fitness value on average. The last operator is "mutation," which promotes diversity in population characteristics. The mutation operator allows for global search of the design space and prevents the algorithm from getting trapped in local minima. Many references discuss details of GAs; for example, see Ref. 20.

Two-Branch Tournament GA

The basic idea in the two-branch tournament GA is that the selection operator evaluates the generated designs against two objective functions independently. This allows for the generation of designs that approximate the Pareto optimal set for a two-objective problem. The major advantage of this method is that it makes use of the large number of individual designs generated by the GA in all generations to find non-dominated designs without significant computational burden beyond what is required for a single-objective GA.[21]

Figure 1 summarizes the mechanics of the two-branch tournament. In a two-branch tournament, individuals are selected to survive as parents by comparing members of the current population based on each objective. By doing this, half of the surviving parents are selected based on their performance on the first objective's fitness function, and half of the parents are selected based on their second objective's fitness function value. The two halves are mixed in the parent pool and then crossed over and mutated to produce the population of the following generation. This process guarantees that parent designs that are better on both objective functions will be selected more often as parent designs to produce offspring designs that have a high probability of performing well on both objectives. This selection mechanism still introduces enough diversity in the population, which is needed to generate designs at the far ends of the Pareto set (see Ref. 21). There are various mechanisms that can be implemented to guarantee the generation of a good approximation of the Pareto front; some of these are elitism and proper selection of the parent-mixing ratio for crossover.[22]

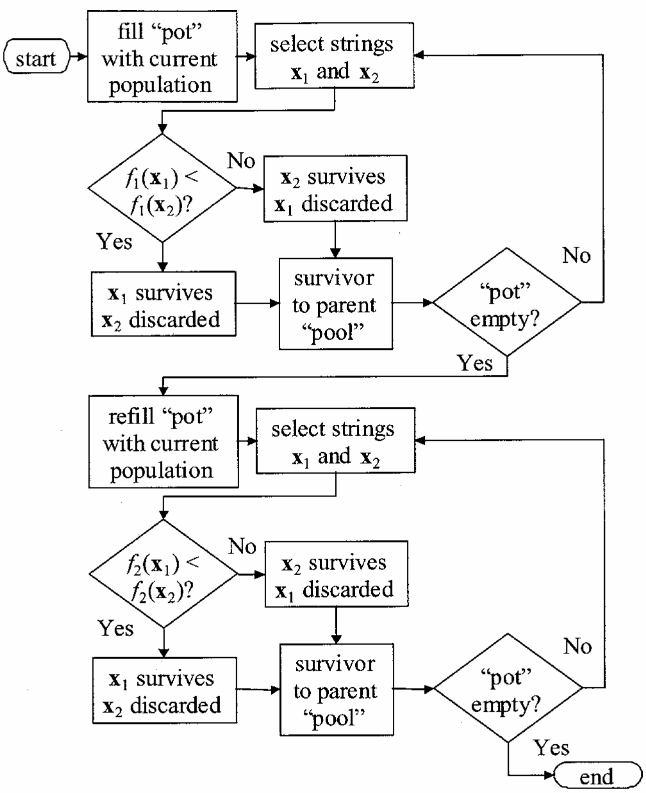

Fig. 1 Two-branch tournament flowchart.

For constrained problems, fitness functions often use exterior penalty functions. Two-branch tournament uses this idea with a slight modification to allow for objectives of different magnitudes. The first fitness function f_1 is computed using Eq. (9). Equation (10) then computes a penalty-scaling factor for the second fitness f_2. The second fitness function is calculated with Eq. (11). Using a penalty-scaling factor in this way penalizes the second objective by the same amount by which the first was penalized:

$$f_1(\mathbf{x}) = \phi_1(\mathbf{x}) + \sum_{j=1}^{J} c_j \max[0, g_j(\mathbf{x})] + \sum_{k=1}^{K} c_k |h_k(\mathbf{x})| \tag{9}$$

$$P^* = f_1/\phi_1 \tag{10}$$

$$f_2 = P^* \phi_2 \tag{11}$$

The penalty multipliers c_j and c_k are selected through experimentation and/or prior experience with the problem. Those multipliers should add penalties for violated constraints of the same order as the objective function. Handling the constraints in this manner discourages the propagation of infeasible designs. A filtering process ensures that only nondominated designs with no violated constraints appear in the approximate Pareto set.

Problem Statement

For the communication satellite design optimization problem, there are two objectives of interest that are shown in Eqs. (12) and (13). The first objective is the minimization of the overall satellite launch mass. This objective was chosen as a representative of the spacecraft cost; it is commonly discussed that the cost of placing 1 kg in Earth orbit is approximately \$30,000 (Ref. 23). The second objective is the maximization of the overall satellite reliability, which is represented in Eq. (13) as a minimization of the risk function of the satellite:

$$\text{minimize}\ \phi_1 = \text{satellite launch mass} \tag{12}$$

$$\text{minimize}\ \phi_2 = \text{satellite risk} = 1 - \text{satellite reliability} \tag{13}$$

Design Parameters

There are 14 design parameters that describe the satellite payload and subsystems and 13 design parameters that describe redundancy options for the satellite subsystems, making a total of 27 design parameters. The first 14 design parameters are summarized in Table 2. In this problem statement, all 14 of these design parameters are discrete parameters; more details about those design parameters may be found in Ref. 11.

Table 2 Satellite sizing tool design parameters[11]

Design parameter no.	Description and discrete values
1–8	HPA type (TWTA or SSPA)
9	Launch vehicle (see Table 1)
10	Solar array cell type (GaAs single junction, GaAs multijunction, Si thin, Si normal, or hybrid Si with GaAs multijunction)
11	Battery cell type (NiCd or NiH_2)
12	N/S thermal coupling (no coupling or coupling)
13	N/S STK thruster technology (xenon plasma, arcjets, bipropellant, and hydrazine monopropellant)
14	E/W STK thruster technology (bipropellant or hydrazine)

Design parameters 15–23 are integer variables that represent redundancy levels in the Ku-band repeater and the eight C-band repeaters. To represent M-out-of-N systems in this problem, M is fixed and N can vary. For example, the first Ku-band repeater in the existing baseline satellite was designed to use 12 active out of 16 available HPAs. In this case, the required number of available amplifiers is $M = 12$, and the number of available amplifiers is $N = 16$. In the optimization problem statement, design parameter 15 describes whether the level of redundancy for the Ku-band repeater will include 0, 1, 2, or 4 additional HPAs, which would make $N = 12$, 13, 14, or 16. The value of M is fixed at 12 for this problem, so there are no redundant amplifiers if $N = 12$.

Design parameters 16–23 represent the redundancy levels associated with each of the eight C-band repeaters in the same manner as the Ku-band repeater redundancy, except with different values for M and N. The first seven C-band repeaters require six active HPAs for each of them. The level of redundancy for each of those seven repeaters is set at 0, 1, 2, or 3 additional HPAs. This level of redundancy means that each of the seven repeaters has $M = 6$ HPAs and $N = 6$, 7, 8, or 9 HPAs. The last C-band repeater requires two active HPAs and its level of redundancy is set at 0 or 1 HPA, meaning that its $M = 2$ HPAs and $N = 2$ or 3 HPAs. Based on these data, the overall number of required active HPAs for all Ku-band and C-band repeaters in the baseline design can be calculated as 56 HPAs [12 Ku-band HPAs + (7 repeaters × 6 + 2) C-band HPAs)]. The 56 required HPAs could be supported by a number of redundant HPAs ranging from 0 and up to a maximum of 26 redundant HPAs [4 Ku-band HPAs + (7 repeaters × 3 + 1) C-band HPAs)].

Design parameters 24–26 describe whether the design includes one or two duplicate sets of the propulsion, ADCS, and the TCR subsystems, respectively. The last design parameter (27) describes the level of redundancy added to the required solar array area. Using the binary coding of the GA, four discrete values of additional area are represented in this problem. This variable can represent the minimum required array area (0% additional area), which corresponds to lowest mass possible, or to use different levels of string redundancies (2, 4, or 6% additional area), which incorporates higher reliability and heavier mass.

Constraints

There are five constraints applied to the problem at hand. The first and second constraint functions ensure that the solar panel length and the radiator panel height computed by the satellite sizing code do not exceed the height of the fairing of the launcher:

$$g_1 = \frac{\text{solar array panel length}}{\text{fairing height}} - 1 \le 0 \tag{14}$$

$$g_2 = \frac{\text{radiator height}}{\text{fairing height}} - 1 \le 0 \tag{15}$$

The maximum allowable lift mass of the launch vehicle is used as a constraint on the satellite total wet mass, which is

$$g_3 = \frac{\text{wet mass}}{\text{maximum lift mass}} - 1 \le 0 \tag{16}$$

Two additional reliability measures are treated as constraints. The reliability of the launcher based on the number of successful launches out of the total number of launches was included because this is generally a specific mission requirement. A minimum reliability of 90% was chosen; this is shown in Eq. (17). Finally, a constraint is imposed so that the overall reliability of the satellite payload is higher than 90% as shown in Eq. (18). This constraint is added because payload reliability, separate from total spacecraft reliability, is generally a requirement for marketing competitiveness to ensure availability:

$$g_4 = 1 - \frac{\text{launcher reliability}}{90\%} \le 0 \tag{17}$$

$$g_5 = 1 - \frac{\text{payload reliability}}{90\%} \le 0 \tag{18}$$

The objective of total spacecraft reliability is the product of the payload reliability and the bus reliability, which includes structure, propulsion, ADCS, TCR, power, and thermal subsystems. Equations (7) and (8) present the calculations for spacecraft reliability.

Results and Discussion

A GA code was prepared using MATLAB® to carry out the multi-objective optimization. The GA employs Gray coding, uniform crossover, and two-branch tournament selection. For this problem, 41 bits encode the 27 design parameters. A constant population size of $41 \times 4 = 164$ individuals and a mutation probability of $(41 + 1)/(2 \times 164 \times 41)$ were used, based on empirical guidelines.[24] To get many nondominated designs for a good approximation of the Pareto front, the GA should be allowed to run for a reasonably large number of generations. The GA was run several times with different starting seeds to determine how many generations are needed to reach a good approximation of the Pareto front. Figure 2 shows that the two-branch tournament returns between 40 and 50 Pareto points as long as the number of generations exceeds 100. Figure 2 shows the number of Pareto points that resulted from 12 different GA runs, each starting with a new random seed.

The two-branch GA was used to generate solutions to the multi-objective satellite design problem. In 300 generations, the two-branch tournament produced 50 points on the Pareto front, representing 50 satellites that tradeoff minimum launch mass and minimum risk (maximum reliability). These 50 designs plotted in a semilog scale appear in Fig. 3. Three regions of the tradeoff space are evident in Fig. 3.

Region 1 contains only two designs with very low reliability and low mass, whereas region 2 contains several designs with a clear

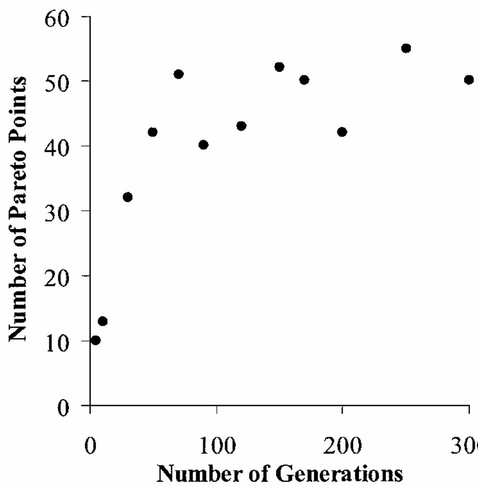

Fig. 2 Performance of the two-branch tournament with the number of generations for 12 GA runs with different random starting seeds.

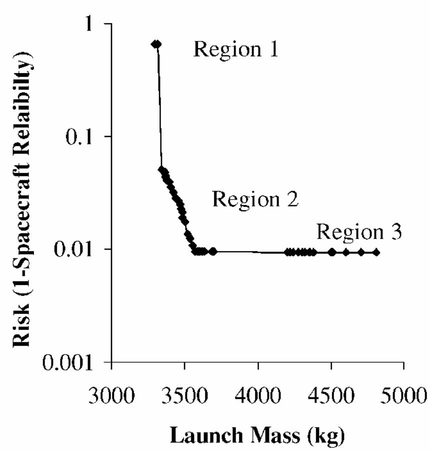

Fig. 3 Pareto optimal set for the satellite design problem using two-branch tournament after 300 GA generations.

tradeoff between the mass and the satellite reliability. In region 2, the reliability values range from 0.9495 to 0.9905, whereas the launch mass values range from 3342.46 to 3693.3 kg. On the other hand, region 3 contains a range of satellite design solutions with approximately same maximum spacecraft reliability value of 0.9906 (minimum risk), but with a wide array of launch mass values that range from 3700 to 4800 kg. To understand the characteristics and the distribution of the design solutions along the Pareto front, the three regions of the Pareto front were analyzed in detail.

Region 1 of the Pareto Front

The two designs in region 1 share the same design variables, except that the higher reliability design contains an additional redundant payload C-band HPA. For this problem, a minimum of 56 HPAs is specified. The low-mass design in region 1 has three redundant HPAs for the C-Band repeaters. The larger design in region 1 has four redundant HPAs, a total launch mass of 3316.52 kg, and an overall reliability of 0.3476. Recall that the original design for this satellite mission has 56 required HPAs and 26 redundant ones (56 out of 82), which represents about 46% redundancy. The payload reliability for the larger design in region 1 is high compared with the overall spacecraft reliability at 0.9827, but the overall reliability is low because of the lack of redundancy in any of the bus subsystems. The two solutions in region 1 have no redundant propulsion subsystem, nor ADCS, nor additional solar array strings. The lack of redundancy in the bus subsystems is also the reason why the mass of those designs is relatively low. The selections for the first 14 design variables of the designs appear in Table 3.

Region 2 of the Pareto Front

In region 2, the GA choices for the types of the technologies used in payload and bus subsystems design are almost the same for all of the solutions obtained; these are also the same choices as the region 1 solutions and are summarized in Table 3. The differences in the design solutions' mass and reliability values are due to the level of redundancy chosen for each subsystem in each solution.

Table 4 shows the levels of redundancies in four sample design solutions from region 2 of the Pareto front shown in Fig. 3. Those sample designs in Table 4 show that, given the same choices for the subsystems technologies, as the level of redundancy increases the system reliability and mass also increase. For the four designs shown in Table 4, the GA chose two sets for the TCR subsystem and the maximum 6% additional solar array area redundancy.

Region 3 of the Pareto Front

The designs in region 3 have more or less the same reliability value, which is the maximum available, with a wide range of mass values. All of the designs in that region have full bus subsystems redundancies, which result in the maximum bus reliability. The redundancy values of the nine repeaters vary slightly from one design to another in this region; the total number of all Ku-band and C-band repeaters' HPAs ranges from 65 to 69. This small variation in payload redundancy explains why all of the designs in region 3 have essentially the same value for the overall spacecraft reliability.

The large spread of the mass values for region 3 designs results from the different types of HPAs and different types of propulsion technologies for both N/S and E/W STK used in these spacecraft designs. Table 5 shows the number of each HPA type used for the eight C-band repeaters and the STK propulsion technologies for five designs selected from region 3 of the Pareto front, along with the launch mass of these spacecraft designs. The satellite reliabilities for the five designs are about the same from a practical standpoint and they range from a low value of 0.9905 for design 5 to a high value of 0.9907 for design 9.

Design 5 has a low mass compared to design 9 for two reasons. First, design 5 has a higher number of repeaters that employ SSPAs, which are lighter in weight than TWTAs. This decreases the overall payload mass, which, in turn, affects the overall spacecraft mass significantly because the spacecraft dry and wet mass are directly proportional to the payload mass. Second, design 5 employs propulsion technologies with much higher I_{sp} for both N/S and E/W STK than design 9.

Table 3 Description of the design solutions parameters in regions 1 and 2

Design parameter	Technology chosen
HPA type	SSPAs for C-band repeaters 1 and 3–8, TWTAs for C-band repeater 2[a]
Launch vehicle	Delta, has the highest reliability index[b]
Solar array cell	Si cells, half weight as GaAs cells, more reliable, but less power output
Battery cells	NiH_2 cells, higher capacity
Thermal coupling	No N/S coupling, more reliable
N/S STK	Plasma propulsion, highest I_{sp}
E/W STK	Bipropellant thrusting, highest I_{sp}

[a]See Ref. 11 for discussion. [b]See Table 1.

Table 4 Subsystems redundancy levels for region 2 of the Pareto front

Design	Mass, kg	Satellite reliability	Number of HPAs	Propulsion sets	ADCS sets
1	3342.46	0.9495	60	1	1
2	3409.39	0.9645	61	1	2
3	3505.32	0.9826	62	2	2
4	3636.16	0.9905	68	2	2

Table 5 Selected design descriptions from region 3 of the Pareto front

		C-band repeater types		STK propulsion	
Design	Mass, kg	SSPAs	TWTAs	N/S	E/W
5	3700.87	5	3	Plasma	Bipropellant
6	4201.56	7	1	Arcjets	Bipropellant
7	4307.17	6	2	Arcjets	Hydrazine
8	4513.74	6	2	Bipropellant	Bipropellant
9	4808.42	2	6	Bipropellant	Hydrazine

Discussion

The satellite sizing tool was integrated with the GA without the reliability model in a previous phase of this research for single-objective optimization to minimize the launch mass of the communication satellite.[11] The addition of the reliability model to the sizing tool increased the number of design variables from 14 to 27, which makes the problem more complex from the computational point of view. However, the choices made by the GA after the addition of the reliability model demonstrate its robustness in handling complex problems. For most of the designs of regions 1 and 2 on the Pareto front, the GA chose SSPAs for all C-band repeaters' HPAs except for the second repeater, whose HPAs were chosen as TWTAs. These are the same results as those generated for the single-objective formulation before including the reliability model. The choice of TWTAs, which are heavier but more power efficient than SSPAs, for the second repeater is due to the high-power amplification required from that repeater.

In the design solutions of regions 1 and 2 of the Pareto front, the GA did not change its choices for the bus subsystems technology options as shown in Table 3. The GA chose the more reliable, less energy generating, and lighter weight silicon cells for the solar array design. The higher energy storing capacity nickel hydrogen cells were chosen for the battery design. Thermal coupling was not applied, which is the same as in the sizing model with no reliability model formulation. Furthermore, plasma propulsion, although less reliable, was chosen for N/S STK because of its significant mass reduction advantage. Finally, bipropellant was chosen for the E/W STK, as was the case in the sizing tool formulation.

The original baseline design for this specific mission required a total of 56 operational HPAs and a total of 26 redundant amplifiers for all Ku-band and C-band repeaters. The results presented in Fig. 3 and Tables 4 and 5 suggest that the original baseline satellite could be considered as an oversized system. Design 4, described in Table 4, uses only 12 redundant HPAs and obtains a payload reliability of 99.999% and a total spacecraft reliability of 99%. Using fewer redundant HPAs reduces the satellite launch mass; this also reduces the satellite acquisition cost because HPAs are generally very expensive.

Some of the GA component and subsystem choices seem nonoptimal at the subsystem level (like the TWTA selection for the second repeater) but result in a lower system wet mass and/or lower system risk as predicted by the sizing and reliability prediction models. Satellite designers might not have considered these somewhat nonintuitive designs without the use of the multi-objective GA. These surprising, but promising, concepts would then become candidates for further design development.

Conclusions

The integration of a multi-objective optimization method, such as the two-branch tournament GA, with system-level satellite design tools has proved to be an effective approach for generating optimal satellite designs in the conceptual design stage. As implemented here for a geostationary communications mission, this approach produced a Pareto optimal set of designs that can help the systems engineers identify major design tradeoffs and evaluate several design concepts across both design objectives. The approach also enforces design constraints and filters the solutions for feasibility to make sure that all designs included in the approximate Pareto set are feasible designs. Providing this type of information about the spacecraft design space should reduce both the time required by and the cost of the spacecraft conceptual design phase.

Analyzing the 50 design solutions generated across the Pareto front for the communications satellite problem provides the satellite designer with a clear understanding of the characteristics of the available design trades. The Pareto front for this problem is divided into three regions, representing different levels of trades between low mass and high reliability. The first region includes designs with low mass and low reliability due to the lack of bus subsystems redundancy. The discrete technology choices incorporated the designs in this first region are selected to minimize the mass of the satellite. Designs in the second region of the front present a more significant tradeoff between mass and reliability. In this second region, the spacecraft designs generally use discrete values of the subsystem technologies that lead to low mass; the variations in reliability among these designs result from the variation of the levels of redundancies for both payload and bus subsystems. The designs found in the third region all have very high-reliability values with a wide range of mass values because of the variation in payload and propulsion subsystem technology choices. These observations are based on the satellite sizing and reliability prediction models developed for this research, but the approach should provide similar insights if alternate sizing and reliability prediction methods were used.

In the satellite dual-objective design problem, the two-branch tournament GA was able generate multiple good designs across the span of the Pareto front in as low as 50 generations with 8200 function evaluations (164 individuals in the population $\times$ 50 generations). The 300-generation run required 49,200 evaluations. With all of the discrete combinations possible in the design space, an enumeration technique would need 2.199×10^{12} function evaluations to cover the whole design space.

In conclusion, the two-branch tournament GA works well for a satellite conceptual design problem, which is posed as a constrained discrete multi-objective problem. Perhaps more important is that using this approach can give very valuable insights into the tradeoffs available in the design space of the system.

References

[1] Gardiner, G., "What's in a Name?," *Manufacturing Engineer*, Vol. 75, No. 2, 1996, pp. 165–169.

[2] Oleson, S., "Influence of Power System Technology on Electric Propulsion Missions," *Proceedings of the Intersociety Energy Conversion Engineering Conference*, Vol. 2, IEEE Publications, Piscataway, NJ, 1994, pp. 750–755.

[3] Meaker, T., "Product Assurance," *Spacecraft Systems Engineering*, edited by P. Fortescue and J. Stark, 2nd ed., Wiley, West Sussex, England, U.K. 1995, pp. 501–528.

[4] Pullen, S., and Parkinson, B., "System Design Under Uncertainty: Evolutionary Optimization of the Gravity Probe-B Spacecraft," NASA CR-20200, CA, Aug. 1996.

[5] Hecht, H., "Reliability Considerations," *Reducing Space Mission Cost*, edited by J. Wertz and W. Larson, 1st ed., Microcosm, Torrance, CA, 1996, pp. 285–300.

[6] George, J., Peterson, J., and Southard, S., "Multidisciplinary Integrated Design Assistant for Spacecraft (MIDAS)," *Proceedings of the AIAA/ASME/ASCE/AHS Structures, Structural Dynamics, and Materials Conference*, Vol. 3, AIAA, Washington, DC, 1995, pp. 1790–1799.

[7] Lamarra, N., and Dunphy, J., "Interactive Sharable Design Environment for Collaborative Spacecraft Design," *Proceedings of IEEE Aerospace Applications Conference*, Vol. 2, IEEE Publications, Piscataway, NJ, 1998, pp. 487–496.

[8] Lamarra, N., and Dunphy, J., "Web-Based Operation of Interactive Sharable Environment for Collaborative Spacecraft Design," *Proceedings of IEEE Aerospace Applications Conference*, Vol. 2, IEEE Publications, Piscataway, NJ, 1999, pp. 43–49.

[9] Fukunaga, A., Chien, S., Mutz, D., Sherwood, R., and Stechert, A., "Automating the Process of Optimization in Spacecraft Design," *Proceedings of IEEE Aerospace Applications Conference*, Vol. 4, IEEE Publications, Piscataway, NJ, 1997, pp. 411–427.

[10] Mosher, T., "Conceptual Spacecraft Design Using Genetic Algorithm Trade Selection Process," *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 200–208.

[11] Hassan, R., and Crossley, W., "Conceptual Design of Communication Satellites with a Genetic Algorithm," *Proceedings of the AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Vol. 5, AIAA, Reston, VA, 2001, pp. 3263–3273.

[12] Pareto, V., "General Notion of Economic Equilibrium," *Manuale di Economia Poltica*, Societa Editrice Libraria, Milan, 1906 (translated to English by A. S. Schwier, *Manual of Political Economy*, Macmillan, New York, 1971, pp. 103–180).

[13] Vincent, T., and Grantham, W., *Optimality in Parametric Systems*, 1st ed., Wiley, New York, 1981, pp. 121–145.

[14] "Multiobjective Optimization," *Matlab Optimization Toolbox User's Guide*, Ver. 2, MathWorks, Natick, MA, July 2002, pp. 3.38–3.97.

[15] Das, I., and Dennis, J., "Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems," *SIAM Journal on Optimization*, Vol. 8, No. 3, 1998, pp. 631–657.

[16] Timmins, A., "A Study of Total Space Life Performance of GSFC Spacecraft," NASA TN-D-8017, MD, July 1970.

[17] Strauss, R., "Reliabilty of SSPA's and TWTA's," *IEEE Transactions on Electron Devices*, Vol. 41, No. 4, 1994, pp. 625–626.

[18] Parkinson, R., "The Hidden Costs of Reliability and Failure in Launch Systems," *Acta Astonautica*, Vol. 44, No. 7, 1999, pp. 419–424.

[19] Isakowitz, S., Hopkins, J. P., and Hopkins, J. B., *International Reference Guide to Space Launch Systems*, 3rd ed., AIAA, Reston, VA, 1999, pp. 1–15.

[20] Goldberg, D., *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed., Addison Wesley Longman, Reading, MA, 1989, pp. 19–63.

[21] Crossley, W., Cook, A., Fanjoy, D., and Venkayya, V., "Using the Two-Branch Tournament Genetic Algorithm for Multiobjective Design," *AIAA Journal*, Vol. 37, No. 2, 1999, pp. 261–267.

[22] Fanjoy, D. W., Beaver, A. K., and Crossley, W. A., "Empirical Studies of Population Size and Parent Mixing for a Multi-Objective Genetic Algorithm," AIAA Paper 2000-4892, Sept. 2000.

[23] Ashley, S., "Bringing Launch Costs down to Earth," *Mechanical Engineering*, Vol. 120, No. 10, 1998, pp. 62–68.

[24] Williams, E., and Crossley, W., "Empirically-Derived Population Size and Mutation Rate Guidelines for a Genetic Algorithm with Uniform Crossover," *Soft Computing in Engineering Design and Manufacturing*, edited by P. Chawdhry, R. Roy, and R. Pant, Springer-Verlag, London, 1998, pp. 163–172.

I. E. Vas
Associate Editor