JOURNAL OF SPACECRAFT AND ROCKETS
Vol. 35, No. 5, September–October 1998

Impact of a Distributed Architecture for Space-Based Radar

Douglas P. Wickert,* Graeme B. Shaw,† and Daniel Hastings‡
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Through a process of system architecture analysis, system cost modeling, and system architecture optimization, the feasibility is assessed of performing the surveillance function of the next generation airborne warning and control system mission from a space-based radar platform. A distributed operations concept is considered that reduces the size of the satellites required, increases system reliability, improves system performance, and reduces the system cost compared to monolithic designs. The system analysis methodology relies on the specification of a set of mission performance requirements. Competing system architectures can then be compared quantitatively on the basis of a simple cost metric. The cost metric used reflects the cost to initial operating capability and consists of subsystem, reliability, and constellation components. The architecture design process minimizes the cost metric while ensuring compliance with the established performance requirements. It is shown that there is an optimal level of distribution that minimizes the cost.

Nomenclature

A_e = antenna effective area, m^2
A_I = footprint area of radar beam, m^2
A_{search} = area of theater to be searched, m^2
a = orbit semimajor axis, m
B = receiver noise bandwidth, Hz
cov_{av} = average level of coverage, number of satellites
cov_{min} = minimum level of coverage, number of satellites
D = antenna diameter, m
dis = discount rate
G = antenna gain
k = Boltzmann constant, 1.38×10^{-23} J/K
k_A = ellipse factor
k_{cov} = constellation coverage gradient
L = total system losses, dB
m = detection integrity parameter
m_s = minimum cluster size for detection
N = constellation size, number of satellites
n = number of integrated pulses
n_s = cluster size, number of satellites
P_{av} = average transmit power, W
P_d = probability of detection
P_{fa} = probability of false alarm
P_t = burst transmit power, W
R_m = mission reliability
R_s = satellite reliability
R_0 = reliability of baseline satellite
r = radar range, km
S/N = signal-to-noise ratio
$(S/N)_1$ = single pulse equivalent signal to noise
T = noise temperature, K
T_d = mean time to detection, s
T_{fa} = mean time between false alarms, s
Δ = dwell time of beam over target, s
η_{int} = integration efficiency
ν = redundancy parameter
σ = radar cross section, m^2
σ_{cov} = standard deviation of coverage, number of satellites
τ = radar pulse width, s
$\phi_A, \phi_P, \alpha_A, \alpha, \beta$ = regressed Unmanned Spacecraft Cost Model parameters
ϕ_d = development cost factor
ϕ_N = constellation cost factor
ϕ_R = reliability cost factor

Introduction

THE development of small, low-cost satellites offers new horizons for space applications when several satellites operate cooperatively. The vision of what can be achieved from space is no longer bound by what an individual satellite can accomplish. Rather the functionality is spread over a number of cooperating satellites. Further, these distributed systems of satellites allow the possibility of selective upgrading as new capabilities become available in satellite technology.

The term *distributed satellite system* can have two different meanings.[1] The first is defined as a system of many satellites that are distributed in space to satisfy a global (nonlocal) demand. Broad coverage requirements necessitate separation of resources. At any time, the system supports only singlefold coverage of a target region. The local demand of each region is served by the single satellite in view. Here, the term distribution refers to the system being made up of many satellites that work together to satisfy a global demand. The second is a system of satellites that gives multifold coverage of target regions. The system, therefore, has more satellites than the minimum necessary to satisfy coverage requirements. A subset of satellites that are instantaneously in view of a common target can be grouped as a cluster. The satellites in the cluster operate together to satisfy the local demand. Note that the cluster may be formed by a group of formation-flying satellites or from any subset of satellites that instantaneously share a common field of regard. The cluster size and orientation may change in time, as a result of orbital dynamics. In any case, the number of satellites in the cluster is bounded by the level of multifold coverage. In this context, distribution refers to the fact that several satellites work together to satisfy a local demand. The entire system satisfies the global demand.

The most important characteristic of all distributed systems, common to both of the defined concepts, is that more than one satellite is used to satisfy the global demand. This is the basic distinction between a distributed system and a singular deployment. Within the classification of a distributed system, the main difference between the two concepts described lies in the way that the local demand is served. Specifically, the distinction is the number of satellites used to satisfy this local demand. The set of satellites that are used to serve the local demand is defined as a cluster. The cluster size, defined as the number of satellites making up the cluster, therefore is a valid measure of the level of distribution. The lowest level of distribution, with a cluster size of one, corresponds to the first meaning of distribution described. Larger cluster sizes correspond to higher levels of distribution.

There are several characteristic features of distributed satellite systems that distinguish them favorably from singular deployments.

Received Nov. 4, 1997; revision received May 5, 1998; accepted for publication May 9, 1998.

*Graduate Assistant, Department of Aeronautics and Astronautics.
†Graduate Fellow, Department of Aeronautics and Astronautics.
‡Professor, Department of Aeronautics and Astronautics. Fellow AIAA.

1) Decentralization of resources means that the systems have intrinsically better survivability and fault tolerance. The distribution in space of the system components also means that revisit times over points on the Earth's surface can be reduced. 2) Smaller, less complex satellites can mean reduced development time and cost. Also the smaller satellites permit the use of low-cost, small commercial launch vehicles. 3) The distributed system can be made adaptable to changes in role or technology. As the requirements or state of the art changes, new system components can be launched to replace the obsolete elements. 4) The reduced range to target of low-Earth-orbiting satellites means that passive remote sensors can support much higher spatial resolution, whereas active devices require less power for operation.

These features make distributed architectures particularly attractive to military applications that require high performance from low-cost, survivable platforms. However, commercial ventures, offering capital gain and being dominated by competition, are certain to exploit the concept earlier and within the very near future. Although it is unlikely that distributed systems are suitable for all applications, there are many missions for which small sensors on many satellites scale very well and give cost-effective solutions. Some applications are distributed implementations of traditional single-satellite deployments, and some represent new capabilities that would be otherwise impossible to achieve.

In this work we focus on the question of the efficacy of using a cluster of satellites to undertake a military space mission. To be specific, we consider the application of space-based radar (SBR) and assess the effect (on cost and effectiveness) of distributing the function of radar detection among the cluster of many satellites.

Overview

The current radar detection of aircraft by the U.S. Air Force is done by the Airborne Warning and Control System (AWACS) platform, which is called the E-3 Sentry.[2§] The E-3 is a modified Boeing 707 airframe with a 30-ft-diam radar dome positioned above the fuselage. The E-3 provides mobile, all-weather surveillance, command, control, and communications (C^3) to both battlefield command centers and pilots. The E-3 can detect and track enemy aircraft and ships, update the location and status of friendly aircraft and ships, support air to ground operations, direct interception of hostile aircraft, and perform the identification friend or foe (IFF) function. The drawbacks of the current E-3 AWACS platform include the time required to respond to a new threat associated with moving assets to the theater, the immense support structure that must also be deployed, technology that is beginning to become outdated, and high year-to-year operating and maintenance costs. It is in each of these areas that a space-based surveillance platform could improve the capability of AWACS. A space-based platform could respond to new threats almost instantaneously, requires relatively little support structure, and is dominated by acquisition and deployment costs rather than by operation and maintenance costs. The challenges to a space-based AWACS platform are a moderate to high level of technological risk and a high cost to initial operating capability.

The AWACS mission functions include surveillance, C^3, and IFF. For a space-based AWACS mission, the space segment of the system will probably only perform the surveillance and possibly the communication components of the AWACS functions. Command and control can be done from a command center linked to the space segment. It is conceivable that the command center will be situated in the continental United States and far removed from the theater of operations. Communications to theater can be done through the radar satellites or separate communications satellites. We concentrate on design and analysis of the space-based surveillance segment of the next generation AWACS mission. Particularly, we concentrate on the search function of four possible surveillance tasks: search, track-while-search, track, and continuous track.

Before any action can be taken against a target, the presence of the target must be determined. This is the purpose of the search mission task. The search process involves searching a given area with a beam footprint that is generally much smaller than the area to be searched. Thus, to cover the entire search area, the beam must be scanned through the area. Using an updated target position with positions from previous search scans to form a track is the track-while-search mission task. Because the track is not sampled at a rate sufficient to predict the target position at the next scan, track-while-search is classified as search radar. Some scenarios require a higher sampled track, such as precision intercept vectoring or handoff to another tracking radar. In continuous track, the radar is used to illuminate the target for some other platform, particularly for intercept by a missile with no transmitter of its own or for intercept by an aircraft that does not wish to reveal its position.

An SBR designed for search can satisfy most of the requirements associated with all four surveillance tasks (search, track-while-search, track, and continuous track). A moderate aperture SBR has a characteristically large beam footprint on the ground (tens to hundreds of kilometers at uhf). The target is, therefore, unlikely to leave the footprint in the time required for intercept, and no track update is required.

The use of SBR has been proposed many times in the past to undertake the mission of aircraft detection. The traditional solution has been to use large satellites in a constellation supporting singlefold global coverage. Such SBRs have always proved unfeasible due to the very high power or large aperture required to detect a return from small aircraft.

In this study, we consider constellations of low-Earth-orbiting spacecraft and assess the effect of distributing the search function over a cluster of many spacecraft. This permits a smaller power–aperture product on each satellite, at the expense of increasing the number of satellites required. The impacts and effects of alternate architectures are estimated quantitatively with the help of a cost metric. The cost metric used for this study is cost to initial operating capability (IOC) cost. A cost model has been developed to calculate the IOC cost metric, incorporating several submodels that are based on a fairly detailed system analysis. A reliability model, a constellation coverage model, and an operations model that reflects system performance have been developed and integrated. All of the model components are combined into a single cost function that reflects the IOC cost metric, relative to a set of performance requirements.

The cost model has been designed so that the most significant architecture options are independent variables of the cost metric. Then by analytic, semianalytic, and numerical solutions, the cost function can be minimized as a function of the desired architecture variables. This process gives the optimal system architecture, according to the IOC cost metric, that meets given system requirements and performance specifications. The optimization process gives optimum values for several architecture variables, including orbital altitude, cluster size, redundancy, satellite power and aperture, and the satellite area search rate. These are the most important architecture variables as they define the concept, and all other system variables may be derived from them. By fixing a set of mission performance requirements, competing system architectures, including distributed and traditional singular deployments, can then be fairly compared on the basis of the IOC cost.

Radar Fundamentals

In this section, we consider the components of a general radar system and introduce the metrics that are conventionally used to measure performance. Mean time to detection and mean time between false alarms are identified as essential performance requirements that must be specified in the design of a search radar system.

All radar systems consists of the same basic elements: transmitter, receiver, and signal processor. The antenna is often used for both transmitting and receiving, though it need not be. If the antenna is used for both functions, the radar is monostatic. A radar with spatially separated transmit and receive antennas is multistatic. Multistatic radars can give very good angular resolution, but at the cost of grating lobes and added complexity due to a need for phase coherence between antennas.

A radar emits electromagnetic (EM) radiation and extracts information from the reflected energy. The vast majority of radars work by analyzing three features of the reflected, received signal:

§Available via the internet at http://www.af.mil/news/factsheets/E_3_Sentry__AWACS.html.

the presence of a reflector, the EM wave's travel time to the reflector, and the frequency content of the signal. The presence of a reflected signal is used to detect targets. The time delay between the transmitted and received signal gives the range to the target if the speed of the EM wave is known for the medium. Finally, the spectrum of the received signal can be used to indicate the velocity of the target relative to the radar by the phenomenon of Doppler shift. The spectrum can also be used to detect moving targets in a stationary clutter background.

Radar Range Equation

The radar range equation relates the maximum range at which targets can be detected to the transmitter power, antenna gain and area, signal-to-noise ratio, signal integration, system losses, thermal noise, and target radar cross section. The signal-to-noise ratio is an important factor in the detection of valid radar returns in the presence of noise. The most general form of the radar range equation is

$$\frac{S}{N} = \frac{P_t G A_e \sigma}{(4\pi)^2 r^4 k T B L} \tag{1}$$

The radar cross section (RCS) of a target is the effective area that reflects power back to the radar receiver. Typical values for the average RCS of common targets range from 40 m^2 for large bomber aircraft to only 0.5 m^2 for conventional winged missiles.[3]

Radar Range Equation for Search Radar

A search radar scans a beam across the search region, detecting returns from potential targets in the beam footprint. Usually, many radar pulses are returned from any particular target on each radar scan, and integration can be used to improve detection. The equivalent signal-to-noise ratio of a single pulse that would give the same detection characteristics as the integration of n separate pulses is given by

$$(S/N)_1 = \eta_{\text{int}} n (S/N) \tag{2}$$

The number of pulses that can be used for integration is bounded by the number of pulses incident upon the target during each scan. This is equal to the product of the pulse repetition frequency (PRF), in hertz, and the dwell time

$$n = \text{PRF} \cdot \Delta \tag{3}$$

Dwell time is defined as the time it takes a point to completely transit the footprint area as the footprint is being scanned and is given by the ratio of the illuminated footprint area A_I to the area search rate (ASR), in square meters per second,

$$\Delta = A_I / \text{ASR} \tag{4}$$

The footprint area illuminated by an equivalent circular aperture with a diffraction limited beam is

$$A_I = [(r/2)(\lambda/D)]^2 \pi k_A \tag{5}$$

The k_A is a function of both orbital radius and frequency. For a fixed target elevation, the elliptical footprint becomes more eccentric as the radar altitude increases because of the increased apparent curvature of the Earth's surface. As frequency increases, the beamwidth decreases, and the footprint elongation due to Earth curvature is less significant. At low elevations, part of the beam extends beyond the limb of the Earth, decreasing the k_A. An average k_A of 4 was determined from simulations and was used in all further calculations.[4]

Replacing the gain G by $4\pi A_e/\lambda^2$ and substituting Eqs. (2–5) into Eq. (1), we obtain the search radar range equation

$$\left(\frac{S}{N}\right)_1 = \frac{P_{\text{av}} A_e \sigma \eta_{\text{int}} k_A \pi}{4^3 r^2 k T B \tau L} \frac{1}{\text{ASR}} \tag{6}$$

where $P_{\text{av}} = P_t(\tau \cdot \text{PRF})$. The search radar range equation depends on range squared instead of range to the fourth and on the power–aperture product instead of the power–aperture–aperture product. A convenient form of the search radar range equation is

$$(S/N)_1 \text{ASR} = K_s P_{\text{av}} A_e \tag{7}$$

where

$$K_s = \frac{\sigma \eta_{\text{int}} k_A \pi}{4^3 r^2 k T B \tau L} \tag{8}$$

Probabilities of Detection or False Alarm

The signal-to-noise ratio is a crucial performance driver for signal processing. The purpose of the radar signal processor is to detect a valid radar signal return in the presence of noise while rejecting false returns. False returns are signals that look like valid targets but are actually due to noise, multipath, clutter, reverberation, or other effects. An envelope detector sets a threshold envelope above the rms noise envelope. The envelope of a signal is the complex amplitude given by quadrature demodulation. When the received signal envelope exceeds the threshold, a detection is recorded. If the threshold is set too low, noise will tend to give too many false detections; too high a threshold results in missing potential targets at moderate signal to noise levels. The probability of detecting a target from the received signal then depends on the signal-to-noise ratio and on the chosen threshold envelope (or equivalently the probability of false alarm). Although there is no closed-form solution, Rice[5] gives a series approximation as for the probability of detection:

$$P_d = \frac{1}{2}\left\{1 - \text{erf}\left[\sqrt{\ln\left(\frac{1}{P_{\text{fa}}}\right)} - \sqrt{\frac{S}{N}}\right]\right\}$$
$$+ \frac{\exp\left\{-\left[\sqrt{\ln(1/P_{\text{fa}})} - \sqrt{S/N}\right]^2\right\}}{4\pi\sqrt{S/N}}\left\{\frac{3}{4} - \frac{\sqrt{\ln(1/P_{\text{fa}})}}{4\sqrt{S/N}}\right.$$
$$\left. + \frac{1 + \left[\sqrt{2\ln(1/P_{\text{fa}})} - \sqrt{2(S/N)}\right]^2}{16\sqrt{S/N}} - \cdots\right\} \tag{9}$$

This relation is plotted in Fig. 1. The signal to noise used in this calculation corresponds to a single pulse. However, Eq. (2) allows us to use the effective signal to noise of n integrated pulses. The calculated probability of detection then corresponds to a complete scan of the radar over the target.

Mean Time to Detection and Mean Time Between False Alarms

Although the probabilities of detection and false alarm are commonly used metrics for assessing the detection performance of radar systems, they can often be misleading. The probability of detection neither carries an associated time bound (integration over different

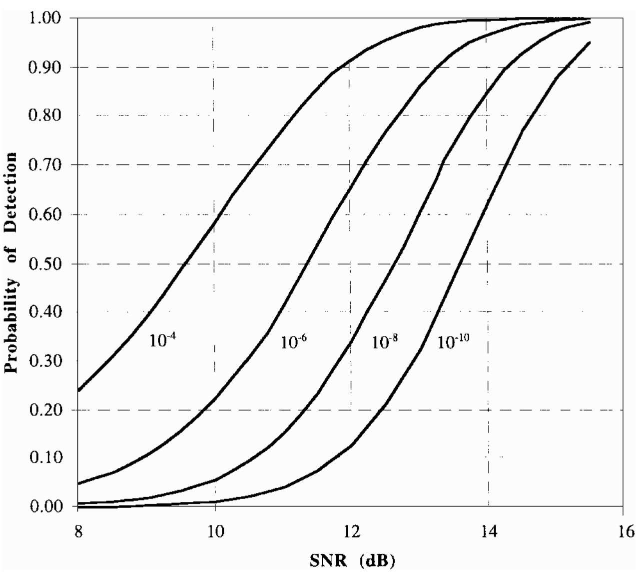

Fig. 1 Probability of detection as a function of signal-to-noise ratio for a range of false alarm probabilities.

dwell times results in very different detection probabilities) nor specifies the area over which the search is performed. A more useful performance metric that captures both the statistical probability of detecting targets and the time required to perform the search is the mean time to detection. This is defined as the average time it takes to detect a target, in the presence of noise, over the entire the region to be searched. Mean time to detection is easily interpreted physically, being naturally related to response time, and leads directly to the establishment of system requirements. Similarly, the mean time between false alarms is more useful than the probability of false alarm.

To derive the mean time to detection for an SBR, the search process can be modeled as two independent negative binomial processes. The negative binomial is a discrete probability distribution that gives the number of trials until the ith successful trial is observed in a process with p probability of success in a single trial. The mean number of trials for i successes in a negative binomial process is $\mu = i/p$.

The first negative binomial process is the number of radar scans past the target until m detections are made. Generally, a detection is not noted unless m of n pulses exceed the threshold. This is known as an m-of-n detection scheme. Because the probability of a single detection in any scan is P_d, the probability of a target declaration during a single scan is P_d/m. In general, the m detections should occur within a sufficiently small window so that some given correlation between scanned positions can be achieved.

The second negative binomial process is the detection of targets throughout the search area. Consider a target randomly located somewhere in the search area. The mean number of trials until the target has been illuminated is

$$\mu_{\text{trials}} = \frac{A_{\text{search}}}{A_I} \tag{10}$$

The mean number of trials until the target has been detected is simply the number of trials to illuminate the target divided by the probability of a target declaration for a single trial:

$$\mu_t = \frac{mA_{\text{search}}}{A_I P_d} \tag{11}$$

The mean time to detect the target is the duration of a single scan, calculated from Eq. (4), times the mean number of trials:

$$T_d = \frac{mA_{\text{search}}}{P_d \text{ASR}} \tag{12}$$

A similar analysis leads to the definition of the mean time between false alarms. With an m-of-n detection scheme, m independent false detections have to occur for each false alarm. The probability of an error (false detection) in a single decision made by the signal processor is P_{fa}. The time between decisions is given by the product of the characteristic pulse width τ and the number of pulses that are integrated between decisions. For a matched filter of bandwidth B, $B\tau = 1$. The mean time to false alarm is, therefore,

$$T_{\text{fa}} \approx mn/BP_{\text{fa}} \tag{13}$$

where n is given by Eq. (3).

Together, the mean time to detection and the mean time between false alarms specify the detection performance and integrity of search radar systems. As such, these two parameters are essential elements in the set of system requirements.

Reasonable Requirements on Mean Time to Detection for SBR

Because mean time to detection is the primary performance metric, it is also the primary system driver. Mean time to detection is the performance specification that scales the entire system. Thus, it is important to establish a good, reasonable requirement for mean time to detection.

Targets must be detected quickly enough to allow adequate time to respond to the threat. The mean time to detection requirement then depends on both the threat and the threat response. The goal of the detection and response process is to react to the threat before the threat can carry out its objective. For example, suppose the threat is a flight of enemy fighter bombers with the objective of bombing an airfield. The targets must be detected soon enough to allow time to decide on an appropriate response, command the response action, and allow the action to take effect before the fighter bombers can make their bombing run. A typical response may be to dispatch fighters or launch surface-to-air missiles (SAMs). If fighters must be scrambled, the response time could be minutes. If SAMs are fired, the response time could be on the order of the missile flight time. The time taken in the command response phase can have wide range. If the scenario is an air battle, the command response is simply the target assignment and may require only seconds. In other scenarios, particularly when the threat is unexpected, the command response may be much longer. Thus, determining reasonable mean time to detection requirements depends on the scenario.

The threat detection and response process is well characterized by the observe, orient, decide, act (OODA)[6] loop. The OODA loop is a major tenet of information warfare (IW) and "describes a single iteration of the cycle proceeding from data acquisition, through information integration and decision making, to enaction of a response."[¶] Operating inside or disrupting the enemy's OODA loop is the primary objective of IW. Target detection is analogous to the observe task, threat processing is the orient task, determine response is the decide task, and the command and response is the act task of the OODA loop.

None of the later steps in the threat detection and response can take place until the targets are detected. Thus, decreasing mean time to detection is important to start the other steps sooner. However, if it requires only seconds to detect targets and minutes to respond to the threat, cutting mean time to detection in half does not significantly reduce the total threat response time while the system cost has likely doubled.

In dialog with personnel at the U.S. Air Force Space and Missile Center, we have established a 10-s mean time to detection as a reasonable requirement for the purpose of sizing of the system. This could actually represent an average mean time to detection over a theater, which includes high-threat regions of 1 s, moderate-threat regions of 10 s, and low-threat areas of 1-min mean time to detection.

There is some flexibility in mean time to detection, even after the system is deployed with a fixed capability. Mean time to detection can be reduced in one region at the expense of an increased mean time to detection somewhere else in the theater. Targets are generally less threatening at longer ranges from friendly forces, resulting in decreased risk for a missed detection. By increasing the area search rate in these low-threat regions, the probability of detection can be reduced to a level commensurate with the risk. This increases the assets available for search in high-threat regions and reduces the mean time to detection for the more threatening targets. The integrity parameter m can be adjusted to give a more optimal search profile. False alarms are less desirable in low-threat regions. Similarly, missed detections are more costly at high-threat regions. Thus, m can be increased while searching low-threat areas and decreased in high-threat areas.

Distributed Architectures for SBR

There have been several previous concepts for distribution of radar, though most of them have been multistatic in nature. NASA, the Air Force Office of Scientific Research, and the Defense Advanced Research Projects Agency all funded investigations of distributed array radar concepts in the early 1980s.[7–9] Most of these concepts used multistatic sparse aperture synthesis from separate platforms to achieve high resolution with limited aperture. In 1996, the U.S. Air Force Space and Missile Center (SMC) initiated a bistatic radar sensor study that examined using uninhabited aerial vehicles as a adjunct receiver to an SBR transmitter.[10] The primary challenge of these and other multistatic radar concepts is the requirement to maintain phase coherence between radar systems. The

¶Whitaker, R., GLOSSARY: The Convoluted Terminology of Information Warfare, 1997, available via the internet at http://www.informtik.umu.se/%7Erwhit/IW.html.

number of required connections between satellites, and hence complexity, increases geometrically with distribution. The distributed radar operations concept developed here is not multistatic, and thus, there is no requirement for phase coherence between systems. All processing between distributed satellites is postdetection processing. Distribution is achieved by sharing the function of aircraft detection among several radars.

Distributed Radar Detection

The form of the equation for mean time to detection and the shape of the curves of detection probability vs signal to noise suggests that there may be advantages in distributing the search function among a cluster of satellites. Equation (12) relates the mean time to detection to the ASR and the probability of detection P_d for a single satellite. To assess the impact of distribution, consider the average detection rate, equal to the inverse of the mean time to detection. The combined average detection rate of a cluster of m_s satellites independently searching the same region is simply the sum of the individual average detection rates of each satellite:

$$\frac{1}{T_d^{(\text{cluster})}} = \sum_{m_s} \frac{1}{T_d^{(\text{sat})}} \tag{14}$$

The component satellites in a cluster can, thus, afford to detect targets at a lower rate, provided their aggregate detection rate (or mean time to detection) satisfies requirements. For an overall mean time to detection requirement of T_d each satellite in a cluster of size m_s only needs to be capable of an individual mean time to detection of $m_s T_d$. Equivalently,

$$m_s P_d \text{ASR} = \frac{m A_{\text{search}}}{T_d} = \text{const} \tag{15}$$

This equation shows that the probability of detection (for each satellite) and ASR (for each satellite) can be traded against cluster size. Reducing the probability of detection requirement for each satellite permits a smaller signal-to-noise ratio that, mated with a smaller ASR, leads to a smaller power aperture product from Eq. (7). The advantage of reducing the probability of detection is to move down the curves of Fig. 1, away from the region of marginal returns at large values of P_d to the linear regime at moderate values of P_d. Similarly, the mean time between false alarms is reduced linearly by adding satellites, so that

$$T_{\text{fa}} = \frac{mn}{m_s B P_{\text{fa}}} \tag{16}$$

Note that the reductions in power and aperture on each satellite are insufficient to offset the increases in total power–aperture resources that results from adding satellites. However, the individual antennas and power systems are much smaller and more feasible for manufacture and launch. In addition, leverage can be gained in several other areas, so that a distributed architecture can actually cost less than singular deployments of comparable performance. The most dominant factor here is an inherent insensitivity to failures, leading to an improved mission reliability.

Mission Reliability and Redundancy

Mission reliability is the cumulative probability that the system has not failed, where failure is defined as a noncompliance of mission requirements. As discussed earlier, the mission requirements for a space-based search radar system are a mean time to detection and a mean time between false alarms. The desired mission reliability is often specified as a system requirement. Because a space-based replacement to AWACS is a critical system for achieving and maintaining air and space superiority, the required mission reliability R_m will be high. For analysis and sizing, we assume an instantaneous 99% mission reliability requirement.

The achievable mission reliability is a function of the individual satellite reliabilities and the system architecture. The individual satellite reliability (R_s) is defined as the probability that a satellite successfully performs its portion of the mission, incorporating the effects of hardware reliability and survivability to hostile actions (jamming, antisatellite devices, etc.). For the simplest case of an architecture featuring a single satellite, the mission reliability is equal to the satellite reliability. The probability of failing the mission is equal to the probability of failure of the single satellite.

Distribution can lead to an improved mission reliability but only if there is redundancy in the design. A distributed architecture with total resources that exactly satisfy the requirements is a serial system and is subject to serial failure modes. A failure in any component will lead to a failure of the system to meet requirements. For a serial system, the mission reliability is the product of the component reliabilities. Because the reliability of a component is always less than unity, the overall mission reliability decreases geometrically with the number of serial components. The distributed operations concept with no redundancy then results in decreased reliability rather than increased reliability as desired. Only by adding redundancy can a distributed architecture take advantage of parallel reliability. System failure of a redundant architecture requires all parallel paths to fail.

We define redundancy as the amount of spare capability that is deployed compared to that necessary to exactly satisfy requirements. If a cluster of size n_s is deployed with a redundancy of ν, the number of satellites m_s that must continue to operate to satisfy the detection and false alarm requirements is given by

$$m_s = \text{ceiling}[n_s/\nu] \tag{17}$$

The ceiling operator is used to force integrality in the number of satellites. An equivalent interpretation of the redundancy parameter is the amount of over specification in the system requirements. It is typical of the defense acquisition process to pad system requirements with a little bit of margin to ensure system performance when there is uncertainty as to what the actual battlefield requirements should be. In this context, and from Eq. (15), if T_d is the original detection performance requirement, then (T_d/ν) is what the contractor is tasked to provide.

The achievable mission reliability of such a configuration is the probability that at least m_s satellites continue to operate:

$$P(\text{sats} \geq m_s) = \sum_{k=m_s}^{n_s} \frac{n_s!}{k!(n_s-k)!} R_s^k (1-R_s)^{n-k} = R_m \tag{18}$$

Equation (18) can be solved iteratively for the value of R_s that satisfies the mission reliability requirement R_m for a given cluster size and level of redundancy. R_s is plotted as a function of cluster size for various levels of redundancy in Fig. 2. The saw-toothed nature of the curves is a result of the incremental nature of m_s. The individual satellite reliability for a singular deployment is of course equal to R_m.

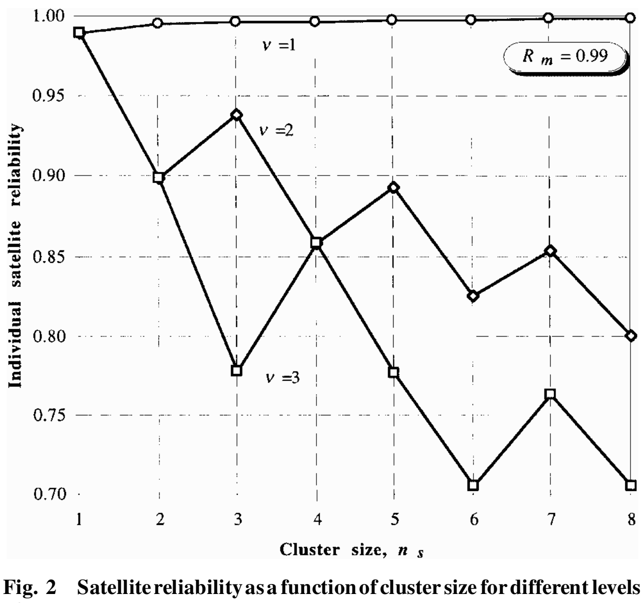

Fig. 2 Satellite reliability as a function of cluster size for different levels of redundancy.

For singularly redundant systems ($\nu = 1$), the satellite reliability increases with cluster size, as expected. For increasing levels of redundancy, a system can afford to lose more satellites and still meet the detection requirement. Therefore, the satellite reliability can be allowed to decrease. For a triple-redundant cluster of size $n_s = 6$, only two satellites are needed to satisfy mission requirements. The system can afford to lose four satellites, and so the satellite reliability need only be 0.71.

Distributed, redundant systems, therefore, need less reliable satellites to achieve the same overall mission reliability. In general, it is substantially cheaper to build satellites of a lower reliability. The cost of reliability can be estimated by considering the additional expenditure that would be needed to improve a satellite from a standard baseline reliability R_0 to the required value R_s, as calculated from Eq. (18). The additional cost to increase a satellite's reliability is expressed as a percentage of the baseline satellite cost.[11] Specifically, if the cost of the baseline satellite is V_0, then the cost of the improved satellite (V_s) is given by

$$V_s = \phi_R V_0 \tag{19}$$

where

$$\phi_R = \frac{\ln(1 - R_s)}{\ln(1 - R_0)} \tag{20}$$

Recall that the desired individual satellite reliability is a strong function of required mission reliability, the cluster size, and the redundancy. The cost factor ϕ_R is, therefore, a function of the same parameters. The cost factor necessary to improve from a baseline reliability of $R_0 = 75\%$ is plotted in Fig. 3 vs cluster size for three different levels of redundancy. Note that, to achieve the required mission reliability of 99% with a single satellite, the cost factor is high, at around 3.4. Compare this to the cost factor close to unity exhibited by a triply redundant cluster with $n_s = 3$. Each of the three satellites in the cluster must be capable of satisfying the detection requirements alone and so need the same power and aperture as the satellite in the singular deployment. Because there are three times more resources on-orbit, one would initially expect the cost to be three times as much. However, as a result of the lower reliability required of each satellite, their individual cost is a third of that for the satellite in the $n_s = 1$ case. As a result, there is no cost penalty for deploying the additional redundancy. Any further cost savings that can be realized from faster movement down the manufacturing learning curve then implies that the overall system cost of the redundant, distributed architectures is less than for singular deployments offering the same performance.

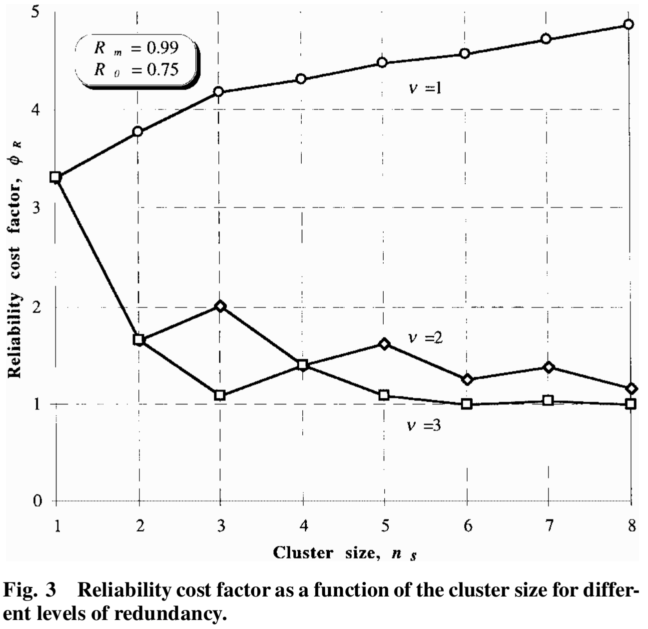

Fig. 3 Reliability cost factor as a function of the cluster size for different levels of redundancy.

Constellation Design and Viewing Characteristics

Constellation size is a significant driver of the system cost; a decrease in the number of satellites required to meet coverage requirements is certainly desirable. The primary goal of constellation design is to optimize the design for given coverage requirements. The chosen constellation should consist of the minimum number of satellites capable of performing the mission, while offering the best coverage possible. In addition, coverage should be concentrated in areas that are likely to be of interest and minimal in regions that are of little interest. Other constellation design considerations are eclipse time, trapped radiation (Van Allen belts), orbit perturbations and subsequent orbit maintenance, orbital debris, and atmospheric drag.

Suitable constellations for the SBR system must be capable of supporting continuous, global coverage of all latitudes between ±75°. The distributed operations concept requires that the chosen constellation gives multifold coverage, equal to the desired cluster size. There is also a constraint imposed on the geometry of the coverage. The elevation angle between the target local horizon and the satellite direction must be between allowable limits. Below a minimum elevation, the transmitted signal is subject to atmospheric ducting and fading. Above the maximum elevation, the target Doppler shift is too low and the clutter return is too high to reasonably discriminate the target from the clutter. For all constellation designs considered, a minimum elevation of 5 deg and maximum elevation of 50 deg is used.

These coverage constraints result in a zenith hole from the perspective of the target. The satellite must be within two concentric cones defined by the constraint elevation central angles. A satellite that is overhead but within the zenith hole formed by the central cone cannot cover the target. From the satellite perspective, the constraint is in the form of a nadir hole.

A great deal of work has been carried out to determine optimal constellations for a distributed SBR system. Because the focus of this paper is not constellation design, only the important results will be presented here. For a more detailed description of the process used, the reader is referred to Ref. 4.

Satellite Coverage and Constellation Size

A heuristic approach was used to develop a set of single-coverage constellations with circular orbits with periods of 1.65, 1.75, 2, 3, 4, and 6 h. Each constellation was designed to give continuous, global access between ±75° latitude, while satisfying the zenith hole constraint. Short coverage gaps on the order of a few minutes exist for some constellations at some latitudes, but this is not a significant problem. For a distributed architecture with a cluster size of n_s, these short coverage gaps simply mean that only $(n_s - 1)$ satellites are visible for the duration of the gap. The impact of these short-lived events is minimal. This set of zenith hole constellations are summarized, along with their coverage statistics in Table 1. The min-mean coverage is the average number of satellites in view at the latitude with the worst coverage statistics. The minimum coverage is the minimum number of satellites in view for any latitude at any time. Note that the average coverage exceeds unity in every case; the constellations were constructed to satisfy minimum coverage requirements.

These heuristically derived constellations, therefore, provide continuous singlefold viewing over all latitudes of interest. However, the distributed operations concept requires continuous multifold coverage equal to the cluster size. Simulations have shown that average coverage scales linearly with constellation size (Fig. 4). A constellation of $2N$ satellites will have an average coverage twice that of a constellation of N satellites. It is tempting, therefore, to simply multiply the zenith hole constellations by the desired cluster size. However, cluster size is defined by the minimum coverage, and this does not scale linearly with constellation size. As the constellation size increases, the coverage becomes more uniform and the minimum coverage for a given latitude approaches the average coverage, in accordance with the central limit theorem. A reasonable conjecture is that minimum coverage actually represents the average coverage less 2 or 3 standard deviations. Two standard deviations will explain over 95% of the observations if the underlying distribution is Gaussian. Another aspect of the central limit theorem is that averaging processes approximate normal distributions for sufficiently

Table 1 Zenith hole constellations

Coverage statistics	A	B	C	D	E	F
Period, h	1.65	1.75	2	3	4	6
Radius, DU	1.11	1.155	1.263	1.655	2.005	2.627
Altitude, km	700	1000	1700	4200	6400	10400
Total satellites	40	33	24	12	9	7
Planes	8, 4	6, 3	6	6	3	7
Satellites per plane	4, 2	4, 3	4	2	3	1
Plane inc., deg	65, 20	67, 30	55	45	45	35
Mean coverage	2.47	2.45	1.77	1.68	1.61	1.51
Coverage standard deviation	0.8935	0.4827	0.2255	0.2805	0.2504	0.2232
Min-mean coverage	1.485	1.741	1.393	1.274	1.380	1.232
Min-mean coverage standard deviation	0.0772	0.0481	0.1109	0.0854	0.0779	0.0250
Latitude of min-mean coverage, °	±75	±75	±75	±75	±50	±75
Minimum coverage	0.84	1.21	0.95	1.15	1.06	1.15
Latitude of minimum coverage, °	±40	±5	±5	±45	±25	±70

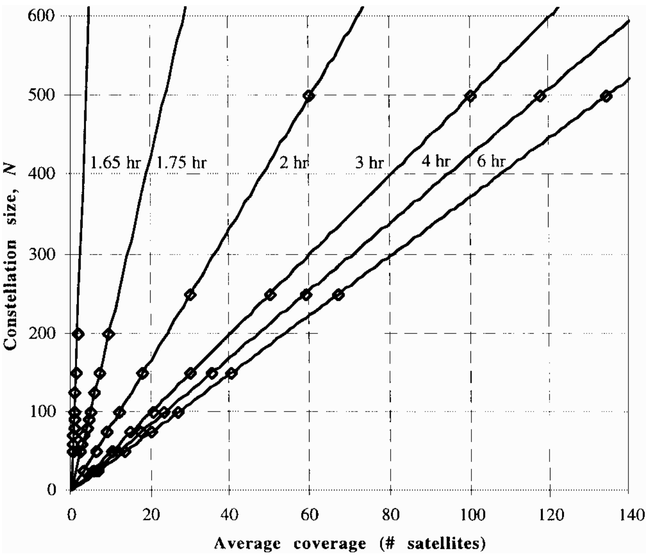

Fig. 4 Constellation size *N* vs average coverage; constellations of arbitrary size were simulated and the average coverage determined from the results.

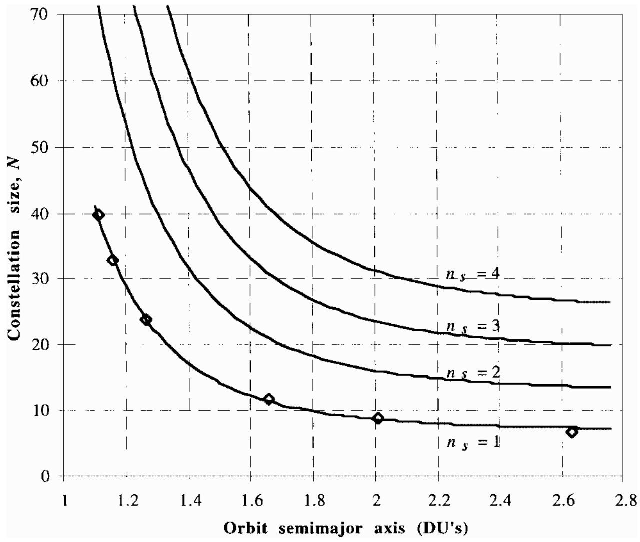

Fig. 5 Constellation size as a function of orbital radius, for several cluster sizes; data points are for the zenith hole constellations.

large samples. Because the coverage statistics are averaged over a diurnal period of 24 h, a Gaussian distribution should approximate the coverage distribution as a function of latitude reasonably well. Thus, the conjectured relationship between average and minimum coverage, as a function of cluster size is

$$\mathrm{cov}_{\min}(n_s) = \mathrm{cov}_{\mathrm{av}}(n_s) - \frac{2\sigma_{\mathrm{cov}}}{\sqrt{n_s}} \tag{21}$$

This conjecture, verified by simulation results, allows us to calculate the number of satellites N needed in a constellation, for a given cluster size n_s and radius a. We proceed by recalling that constellation size scales linearly with average coverage:

$$N(a, n_s) = k_{\mathrm{cov}}(a)\mathrm{cov}_{\mathrm{av}}(n_s) \tag{22}$$

The slope k_{cov} is independent of coverage and can be determined for each altitude from the single-coverage zenith hole constellations of Table 1. Substituting for $\mathrm{cov}_{\mathrm{av}}$ from Eq. (21) and noting that $\mathrm{cov}_{\min}$ is equal to the cluster size, we obtain a relationship for the constellation size in terms of the cluster size and altitude:

$$N(a, n_s) = k_{\mathrm{cov}}(a)\left(n_s + \frac{2\sigma_{\mathrm{cov}}}{\sqrt{n_s}}\right) \tag{23}$$

To be conservative, this calculation is performed using statistics corresponding to the worst-case latitude. This relation is plotted in Fig. 5 using the zenith hole parameters of Table 1.

Range to Target

A similar analysis methodology was used to determine the maximum slant range to targets from the satellites in a given constellation. Once again employing the central limit theorem, the maximum slant range is taken to be two standard deviations from the average value. Average range and the standard deviations were determined assuming an even distribution of targets throughout the coverage region of a single satellite. Extending to n_s satellites with the help of the central limit theorem, the 2-σ maximum slant range to the target can be calculated and is plotted vs orbital radius in Fig. 6.

Impact of Distribution

Distribution improves constellation coverage by decreasing the variability of coverage over time for a given latitude. In a similar manner, distribution decreases the variability of target slant range, which results in a more efficient design with less excess capability. The distributed radar system also has the advantage of an increased number of perspectives of the target area, which results in improved target detection and clutter rejection capability. Different perspectives of the same target also improves the spatial resolution of its derived location.

System Optimization for the AWACS Mission

The goal of the system optimization procedure is to determine values for important system architecture variables to optimally satisfy the AWACS mission requirements. Optimality is defined as the lowest cost alternative that satisfies the mission requirements. Before describing the optimization procedure, therefore, we define the mission requirements and the cost metric used to measure optimality.

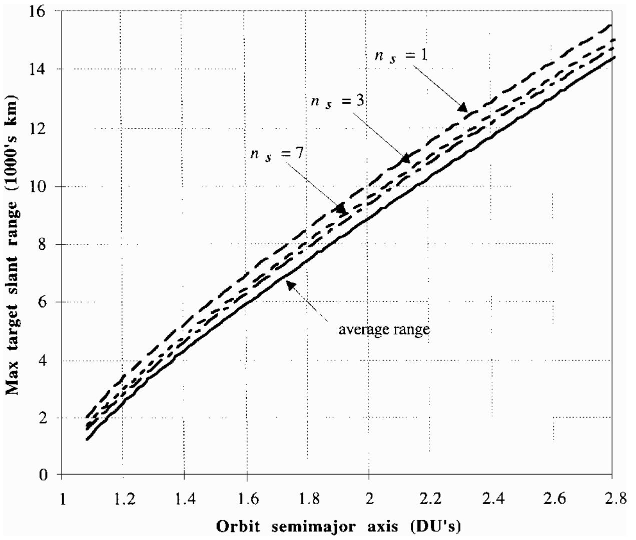

Fig. 6 The 2σ maximum slant range to target as a function of orbital radius.

AWACS Mission Requirements

The mission requirements specify the minimum acceptable levels of performance for a system designed to carry out the AWACS mission. As described earlier, the performance of an SBR system can be measured using the mean time to detection and the mean time between false alarms. Acceptable values of these two parameters must be specified as performance requirements, along with the RCS of typical military targets that must be detected and the size of the area to be searched. The nonzero probability of component failures or hostile actions intended to hinder operations leads to a mission reliability requirement. The system must be able to satisfy performance requirements at the end of life, with a specified probability greater than or equal to R_m, even in the presence of failures or enemy action. Of course, this means that at the beginning of life, or in the absence of failures, the system will satisfy the performance requirements with a probability higher than R_m. The complete mission requirement set is, therefore, 1) mean time to detection $T_d = 10$ s, 2) mean time between false alarms $T_{fa} = 2.75$ h, 3) RCS of targets $\sigma = 1$ m^2, 4) search area $A_{search} = 10^{12}$ m^2, and 5) mission reliability $R_m = 0.99$ over a 10-year lifetime. All systems considered viable alternatives must be capable of satisfying these requirements.

IOC Cost Metric

The cost metric used to measure optimality is the cost to IOC. IOC costs include all development and acquisition costs up to and including the initial deployment costs. Thus, operations costs, support costs, ground facility costs, etc., are not included in the metric. Although operating costs are important for life-cycle system analysis, the baseline missions considered are assumed to have very similar operating costs. Operations costs in the metric would decrease the effect that different system parameters have on the overall cost and disguise the effects of system architecture. Of course, if different architectures have different operational requirements and cost, they must be considered during the comparison and modeling. For the SBR mission, cost of operations are assumed constant across the range of architectures.

The IOC cost is the cost to develop, acquire, and deploy the constellation of satellites. It is convenient to model the cost of each satellite as a function of the most significant subsystem cost drivers, which are in turn functions of the system architecture and operations. The metric is then the product of the cost per satellite and the number of satellites in the constellation N:

$$\begin{aligned}\text{IOC cost} &= \phi_N\, \text{cost}_{\text{sat}} \\ &= \phi_N(\text{cost}_{\text{power}} + \text{cost}_{\text{aperture}} + \text{cost}_{\text{bus}} + \text{cost}_{\text{launch}}) \end{aligned} \tag{24}$$

where ϕ_N is a cost multiplier representing the constellation size N, modified to account for learning curve effects,

$$\phi_N = N(a, n_s)^{1 + \ln(1-\text{dis})/\ln(2)} \tag{25}$$

The discount factor dis used in this calculation is chosen to reflect cost savings from mass production. Large learning curves have not yet been proven in the satellite industry but have been realized in several other industries including automobiles and computers. Until recently, satellites have been built in small numbers and with little standardization between satellites for different missions. The commercial satellite industry, most notably Hughes and Lockheed Martin, are changing this historical trend. Several current concepts should help to validate the learning curve model in the next couple years, notably, the large telecommunications constellations, including Iridium and Teledesic. Lockheed Martin is reportedly realizing 15% learning curves in the production of Iridium buses.

The four components of satellite cost are power, aperture, bus, and launch. These are the significant drivers in development, production, and deployment for the SBR system. Power and aperture costs are very strong functions of architecture. Launch is also a significant function of architecture through constellation size (number of launches) and altitude (size of the booster). Bus costs are included to capture the cost of all remaining subsystems.

Each subsystem cost is calculated using the theoretical first unit cost (TFU) of the U.S. Air Force Unmanned Spacecraft Cost Model (USCM).[12–14] TFU costs are recurring costs and quantify the production phase of the life cycle. Accounting for additional reliability expenditure, we calculate the subsystem production cost as (ϕ_R TFU), where ϕ_R is described by Eqs. (19) and (20). Subsystems developed for satellites placed in orbits within the Van Allen radiation belts are modeled to have higher recurring costs to account for the extra mass needed for radiation hardening. Research, development, test, and evaluation costs are modeled as a constant multiplier of the TFU. A rule of thumb is that development costs typically run 2–3 times the TFU. Each of the subsystem costs is, therefore, calculated as

$$\text{cost}_{\text{subsystem}} = (\phi_R + \phi_d) \cdot \text{TFU}_{\text{subsystem}} \tag{26}$$

where ϕ_d is a development cost multiplier with a (conservative) nominal value of 4.

The cost metric is also a function of several additional parameters, e.g., aperture area density, power density, battery capacity, solar array efficiency, transmit/receive (T/R) module mass, etc. Nominal values were selected to reflect current technology, technology that is expected in the near future, and technology that may be possible in several decades.

Architecture Optimization Procedure

The optimization procedure determines the values of the most important architecture parameters to minimize the IOC cost. Some system parameters (m, PRF, λ, B, L) can be simply assigned reasonable values, either because of technological limitations or because they do not have a significant impact on the overall architecture. These are referred to as the assigned variables. After specifying the system requirements, the remaining undetermined system variables are 1) cluster size n_s, 2) redundancy ν, 3) orbital semimajor axis a, 4) power per satellite P_{av}, 5) aperture per satellite A_e, and 6) ASR. All other parameters can be expressed in terms of these basic quantities. The architecture optimization procedure, therefore, determines the values of these basic parameters that minimizes the IOC cost. Power, aperture, and ASR can be optimized analytically for given values of cluster size, redundancy, and altitude. The optimal n_s, ν, and a are then determined by searching through the parameter space for the minimum cost option.

Power and aperture impact the system cost through the payload cost of each satellite:

$$\begin{aligned}\text{cost}_{\text{payload}} &= \text{cost}_{\text{power}} + \text{cost}_{\text{aperture}} \\ &= (\phi_R + \phi_d)(\text{TFU}_{\text{power}} + \text{TFU}_{\text{aperture}}) \end{aligned} \tag{27}$$

The USCM actually models TFU costs as power laws, regressed from historical data. The payload cost can, thus, be written

$$\text{cost}_{\text{payload}} = (\phi_R + \phi_d)\left(\phi_A A_e^{\alpha_A} + \phi_P P_{av}^{2\alpha_P} + \beta\right) \tag{28}$$

The power and aperture should be chosen to minimize this cost. However, the values of P_{av} and A_e cannot be chosen freely. For the search radar mission, the power–aperture product is constrained to satisfy the radar range Eq. (7). We can use this constraint, with the derivative of the payload cost equation, to optimally allocate expenditure (and resources) between power and aperture, in terms of the other variables of the radar range equation. If the payload cost scaled equally with power and aperture, a hyperbolic constraint such as Eq. (7) would lead to a solution that shared expenditure equally between P_{av} and A_e. The actual TFU scalings result in the following power and aperture optimality equations:

$$P_{av} = \left(\frac{\alpha_A \phi_A}{2\alpha_P \phi_P}\right)^{1/(2\alpha_P + \alpha_A)} \left[\frac{\mathrm{ASR}(S/N)_1}{K_s}\right]^{\alpha_A/(2\alpha_P + \alpha_A)} \tag{29}$$

$$A_e = \left(\frac{2\alpha_P \phi_P}{\alpha_A \phi_A}\right)^{1/(2\alpha_P + \alpha_A)} \left[\frac{\mathrm{ASR}(S/N)_1}{K_s}\right]^{2\alpha_P/(2\alpha_P + \alpha_A)} \tag{30}$$

These conditions represent the optimal allocation of power and aperture for a given value of their product (or equivalently for given values of ASR and signal to noise). We can further optimize by requiring that the power–aperture product be minimized, while still satisfying the detection requirements.

Consider Eqs. (9), (15), (16), and (30), representing the mean time to detection and false alarm requirements, the relationship between detection probability and signal to noise, and the optimality condition on aperture. By solving this system of equations, we can obtain an expression for ASR in terms of only $(S/N)_1$ that is independent of the radar range equation. Substituting into Eq. (7) then gives the power–aperture product as a function of the single variable $(S/N)_1$. Setting the derivative to zero and solving gives the value of $(S/N)_1$ for the minimum power–aperture product that still satisfies mission requirements. The system is fully determined, and optimal values for ASR, P_{av}, and A_e can be calculated. This optimization procedure can be carried out for different values of n_s, ν, and a and the minimum cost architecture can be found.

Optimization Results

The results of the system optimization presented in this section were obtained using the following values for the assigned system parameters: 1) wavelength $\lambda = 10$ cm, 2) receiver bandwidth $B = 1$ MHz, 3) system losses $L = 6$ dB, 4) PRF = 10 kHz, and 5) integrity parameter $m = 1$. No attempt was made to optimize the system with respect to these parameters, and their values were chosen based on other system drivers, current technology limitations or the availability of unclassified data. The choice of wavelength for a real system is complicated by target RCS values and spectrum availability, but for the purposes of initial sizing, a wavelength of 10 cm is considered reasonable. The values of B and PRF are similar to the values used in other SBR studies.[10] Although the absolute value of the IOC metric is affected by these parameters, the important trends are largely invariant of the values. A production learning discount factor of 15% was assumed to be reasonable, given the size of the constellations considered.

Figure 7 shows the IOC cost vs orbital altitude for various cluster sizes, with dual redundancy. As expected, orbital altitude has a large impact on the IOC cost of an SBR system. At high altitudes, the power and aperture is driven to very large values as a result of the range squared scaling of the radar range equation. At lower altitudes, the cost savings arising from a reduction in power and aperture, coupled with an increased leverage from production learning, more than makes up for the corresponding increase in constellation size. This trend is shown clearly in Fig. 7. The minimum cost altitude occurs for all cluster sizes at an orbital radius of 1.11 distance units [DU; one Earth radius (6378 km)]. At altitudes lower than this, the enormous number of satellites and the propulsion systems required to overcome atmospheric drag dominate the cost. The maxima in the curves occurs at an altitude in the Van Allen radiation belts and is a result of the extra mass and development expenditure needed for radiation hardening.

Figure 8 shows the IOC cost as a function of cluster size and redundancy, for the minimum cost orbital radius of 1.11 DU. There are several important trends displayed in Fig. 8. 1) Without redundancy

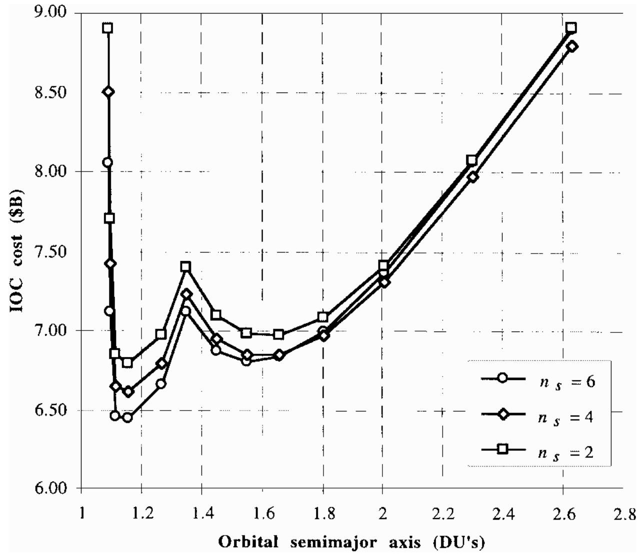

Fig. 7 IOC cost vs orbital radius for three different cluster sizes, with dual redundancy.

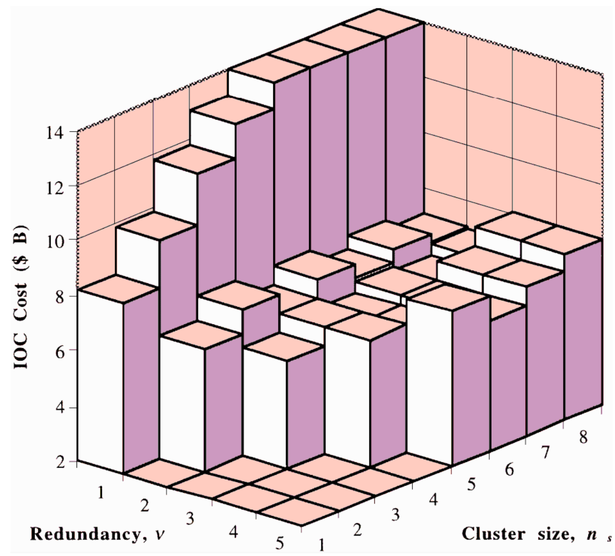

Fig. 8 IOC cost vs cluster size and redundancy, for an orbit with radius 1.11 DU.

($\nu = 1$), the cost increases geometrically with n_s as a result of the increased reliability required of the satellites. 2) The IOC cost has a minimum at a cluster size greater than unity. This is true for all learning curve discount factors less than unity. 3) Small clusters ($n_s \leq 3$) benefit predominantly from increased redundancy. The cost reductions in power and aperture from sharing functionality among a small number of satellites is not as significant to IOC cost as the savings from lower reliability expenditure. Consequently, the minimum cost architecture for small clusters has $\nu = n_s$ (full redundancy). 4) Large clusters ($n_s \geq 4$) benefit predominantly from the power and aperture reductions arising from distributed functionality. The savings resulting from a considerably smaller power and aperture are greater than those from low satellite reliability requirements. High levels of redundancy in large clusters leads to cost penalties because the constellation size is large, but the satellites still need big payloads to perform the mission. Although high levels of redundancy lead to lower satellite reliability requirements, there is little or no cost benefit from increasing beyond triple redundancy (Fig. 3). As a result, the minimum cost architecture for large clusters has only partial redundancy, $\nu \leq n_s$.

Distribution in the forms of sharing the search task among the cluster and deploying spare capability are the most significant contributions to system optimality. Optimal distribution is a function of orbital altitude and learning curve discount. The balance between the number of satellites sharing the task and the number of spare satellites is important. Except for the case of no redundancy, some distribution is always advantageous. The factors that drive cost as a function of distribution are 1) reliability cost decreases with increasing redundancy, 2) decreased power–aperture product per satellites with increased task sharing, 3) number of satellites and launch costs increase with distribution, 4) increased coverage and increased probability for favorable target viewing angles with increased distribution, 5) design efficiency increases with distribution due to decrease in target range variability, and 6) marginal decrease in production learning curves with increasing distribution.

Some of the advantages due to distribution are obvious, e.g., increased reliability and decreased power–aperture product per satellite. There are also several subtle advantages that nonetheless contribute to the reduction in system cost. Variability of range to target decreases with increased distribution, which results in a more efficient design. Coverage efficiency improves in a similar manner. For the same minimum coverage, constellations with higher levels of distribution require slightly smaller constellations than implied by linear scaling. Distribution also increases the target perspectives, which improves performance in clutter.

From these results, candidate architectures can be chosen that can optimally perform the AWACS search radar function. For illustration purposes, we select two alternate architectures: SSL 1, featuring a small cluster size, and SSL 2, which has a larger cluster size. The details of these systems are given in Table 2. Both architectures satisfy the mission requirements equally well. SSL 1 is the lowest-cost alternative but with an rf power requirement of 12 kW may be just beyond the reach of current technology. SSL 2 comes with a higher price tag but has more moderate power and aperture needs ($P_{\text{av}} = 8$ kW and $A_e = 62$ m^2). Both systems benefit from requiring a low satellite reliability of 78%, while still satisfying the mission reliability of 99%. Although this satellite reliability seems very low compared to conventional satellite designs, recall that this value is functionally defined, representing the probability of satellite operation under all conditions, including any enemy's hostile actions. The low reliability of the satellites in SSL 1 and SSL 2 corresponds to the lower level of defensive capabilities that must be equipped on the satellites to satisfy mission objectives.

Table 2 Optimal system architecture summary for SSL 1 and SSL 2

Parameters	SSL 1	SSL 2
IOC cost, \$ × 10^9	6.0	6.3
Cluster size n_s	3	5
Constellation size N	107	176
Redundancy ν	3	3
Satellite reliability R_s	0.78	0.78
Altitude, km	700	700
Semimajor axis a, DU	1.11	1.11
Power P_{av}, kW	12	8
Aperture A_e, m^2	92	61
$(S/N)_1$, dB	14	13.9
ASR, km^2/s	1.2×10^5	0.6×10^5
Probability of detection P_d	0.84	0.84
Probability of false alarm P_{fa}	1×10^{-8}	1.35×10^{-8}
Number of integrated pulses n	100	270
Dry mass, kg	3700	2500
Wet mass, kg	4700	3200

Comparative Analysis

One of the powerful uses of the system cost model is the quantitative comparison on the same basis of different concepts. The optimal systems SSL 1 and SSL 2 are compared with four other SBR concepts that have roughly similar missions in Table 3. SMC 1 and SMC 2 are the proposals from the Space and Missile Center (SMC) Space Sensor Study for the 1995 Corona conference.[10] Iridium is the 66-satellite constellation that will provide global cellular telephone service and is included only for cost model validation. Iridium contracted Motorola to procure and deploy the entire constellation for approximately \$3.5 × 10^9. The SBR cost model predicts \$2 × 10^9 IOC cost for an Iridium-type system. Although accuracy in the absolute cost predicted by the model was never desired, the model gives reasonable results. Nevertheless, the IOC cost model should only be used to compare systems on a relative basis.

SSL 1 and SSL 2 are significantly (5–10 times) cheaper than the SMC concepts. This is in part due to the relaxed mean time to detection requirement. The SMC concepts were sized based on a search area equal to six AWACS horizon circles, updated uniformly every 10 s. They also assumed a minimum–maximum grazing angles of 10/60 deg (as opposed to the 5/50 deg used in the SSL designs). As a result, the SMC systems are over capable compared to the SSL concepts. To fairly compare concepts, the SMC architectures were resized for the same mean time to detection and grazing angles as used in the SSL designs. While maintaining the same altitudes as the original SMC designs, the resized SMC architectures were also optimized for power, aperture, and ASR in the same way as the SSL systems. After this resizing and optimization, the SMC concepts are between 50% and 100% more expensive than the SSL distributed designs for the same level of performance. This shows the potential benefits gained from moving to a distributed architecture.

Conclusions

The initial question that was asked was whether it was advantageous to distribute the function of aircraft detection from space across a number of spacecraft so that the each target area would have multifold coverage. This question was addressed through a detailed process that optimized the system design using the cost to initial operating capability as a fundamental metric.

The answer to this question is that some level of distribution above the minimum required for global coverage is always advantageous in minimizing cost. Distribution allows the task of searching to be shared among several satellites viewing the same area, reducing the resources required of each satellite. In addition, distribution allows spare capability to be deployed. The optimal amount of distribution is a strong function of the redundancy. If no redundancy is employed, all of the satellites must work in series to detect a target within the required time. This forces the satellite reliability to very high values, driving up the system cost. If some redundancy is deployed, then the optimal amount of distribution for each region to be covered is between three and five. This means that each region is being searched simultaneously by three or five satellites. The minimum exists because of a tradeoff between competing trends; larger cluster sizes reduce the power and aperture expenditure for each satellite, whereas smaller clusters require fewer satellites. For very large clusters, the number of satellites in the constellation and the cost associated with their construction and launch dominates the IOC cost. For very small clusters, the power and aperture on each satellite must be very large, driving the system cost away from the optimum.

Table 3 SBR concepts system and performance comparison

System	Alt, km	P, h	P, kW	A, m^2	T_d, s	ASR, km^2/s	Constellation size	IOC cost, \$ × 10^9
SMC 1	2,600	2.35	25	2,275	4.1	2.6×10^5	26	61
SMC 2	10,000	6	30	12,300	3.6	2.9×10^5	9	34
SMC 1 resized	2,600	2.35	30	227	10	1.2×10^5	17	9.2
SMC 2 resized	10,000	6	80	590	10	1.2×10^5	8	12.1
SSL 1	700	1.65	12	92	10	1.2×10^5	107	6.0
SSL 2	700	1.65	8	61	10	0.6×10^5	176	6.3
Iridium[a]	775	1.675	0.9	6	Contract price \$3.5 × 10^9		66	2

[a]Shown for comparison of the cost model output.

The optimum level of redundancy is strongly dependent on the cluster size. Large clusters benefit mostly from a reduction in the power and aperture resources on each satellite. Employing more than dual redundancy to large clusters diminishes this effect while adding more satellites, resulting in a cost penalty. Conversely, small clusters cannot gain enough leverage from sharing the search function between only a few satellites. As a result, they benefit mostly from adding redundancy, because the cost of each satellite is greatly reduced.

Acknowledgments

This work was sponsored by Lockheed Martin. The authors would like to acknowledge guidance and help from Charles McQuiddy of Lockheed Martin as well as many useful discussions with Robert Preston of the Air Force and Vincent Vitto of Massachusetts Institute of Technology Lincoln Laboratories.

References

[1]Shaw, G. B., and Hastings, D. E., "A Generalized Analysis Methodology for Distributed Satellite Systems," International Astronautical Federation International Workshop on Mission Design and Implementation of Satellite Constellations, Toulouse, France, Nov. 1997.

[2]"E-3 Sentry (AWACS)," Air Combat Command Public Affairs Office, U.S. Air Force Fact Sheet 96-13, Langley AFB, VA, 1996.

[3]Cantafio, L. J. (ed.), *SBR Handbook*, Artech House, Norwood, MA, 1989.

[4]Wickert, D. P., "Space Based Radar—System Architecture Design and Optimization for a Space Based Replacement to AWACS," M.S. Thesis, Dept. of Aeronautics and Astronautics, Massachusetts Inst. of Technology, Cambridge, MA, June 1997.

[5]Rice, S. O., "Mathematical Analysis of Random Noise," *Bell System Technical Journal*, Vol. 23, 1944, pp. 282–332; also Vol. 24, 1945, pp. 46–156.

[6]Fadok, D. S., Boyd, J., and Warden, J., *Air Power's Quest for Strategic Paralysis*, Air Univ. Press, Maxwell AFB, AL, Feb. 1995.

[7]Heimiller, R. C., Belyea, J. E., and Tomlinson, P. G., "Distributed Array Radar," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-19, No. 6, 1983, pp. 831–839.

[8]Steinberg, B. D., and Yadin, E., "Distributed Airborne Array Concepts," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-18, No. 2, 1982, pp. 219–226.

[9]"Distributed Array Radar (DAR)," NASA Study, Feb. 1981.

[10]DeLap, R., and Suhr, S., "Transition of Airborne Surveillance/Reconnaissanceto Space," AIAA Paper 96-4411, 1996.

[11]Hecht, H., "Reliability During Space Mission Concept Exploration," *Space Mission Analysis and Design*, 2nd ed., edited by W. Larson and J. Wertz, Microcosm, 1993, pp. 700–714.

[12]*Space Division Unmanned Spacecraft Cost Model*, 5th ed., U.S. Air Force Space Div., Directorate of Cost Analysis, El Segundo, CA, 1981.

[13]*Space Division Unmanned Spacecraft Cost Model*, 6th ed., U.S. Air Force Space Div., Directorate of Cost Analysis, El Segundo, CA, 1988.

[14]Wong, R., "Cost Modeling," *Space Mission Analysis and Design*, edited by W. Larson and J. Wertz, Microcosm, 1992, pp. 715–740.

A. C. Tribble
Associate Editor