

On Mutual Information for Observation-to-Observation Association

Islam I. Hussein, Christopher W. T. Roscoe,
and Matthew P. Wilkins
Applied Defense Solutions, Inc.
Columbia, Maryland 21044
Email: IHussein@AppliedDefense.com

Paul W. Schumacher, Jr.
Air Force Research Laboratory
Kihei, Hawaii 96753

Abstract—In this paper, we build on recent work to further investigate the use of mutual information to tackle the observation-to-observation association (OTOA) problem where we are given a set of observations at different time instances and wish to determine which of these observations were generated by the same RSO. The approach relies on using an appropriate initial orbit determination (IOD) method in addition to the notion of mutual information within an unscented transform framework. We assert that, because the underlying initial orbit determination algorithm is deterministic, we can introduce an approximate correction factor to the IOD methodology. The correction is a function represented by a constant bias for this work but can be expanded to other parametrizations. The application of this correction results in an order of magnitude improvement in performance for our mutual information data association technique over the previous results. The correction can be used in conjunction with other information theoretic discriminators for data association; however it was found in previous work that mutual information is the most precise discriminator. The information theoretic solution described in this paper can be adjusted to address the other (OTTA and TTTA) association problems, which will be the focus of future research. We will demonstrate the main result in simulation for LEO, MEO, GTO, and GEO orbit regimes to show general applicability.

I. INTRODUCTION

Consider a set of indistinguishable objects moving continuously under the influence of a common set of deterministic dynamics and stochastic environmental influences. One or more of these objects appear randomly in the field of view (FOV) of one or more sensors (i.e. they are detectable above the background sensor noise). While these objects persist in the sensor FOV and remain detectable, the sensor provides a set of noisy measurements of the object states and their time stamp, which typically includes a subset of false detections and clutter. The essence of the multi-object tracking problem is to find tracks from these noisy sensor measurements and to rule out clutter from resident space objects (RSOs). The literature is replete with techniques on state estimation if the sequence of measurements associated with each object is known. However, the association between measurement observations and objects is not always known, leading to the well known problem of uncorrelated tracks (UCTs) when attempting to update the

space catalog of observed RSOs. The crux of modern space surveillance from an algorithmic point of view is to solve the data association problem and determine which measurements were generated by which objects.

In general, there are three types of data association problems. The first is the observation-to-track association (OTTA) problem described above, where the analyst seeks to associate each observation with a unique track (or none) given an observation with some known measurement statistics and a set of existing candidate (uncertain) resident space object (RSO) tracks. The second association problem is where we have multiple tracks at different time instances from one or more sensors and wish to determine whether any of the tracks belong to the same RSO. This is the track-to-track association (TTTA) problem. The final association problem is where we are given a set of observations at different time instances and wish to determine which of these observations were generated by the same RSO. This is the observation-to-observation association (OTOA) problem.

In this paper, we continue our work considering the problem of OTOA. This problem can be thought of as determining the statistical dependence of observations, from which there are numerous metrics to choose. Ideally, the measure of statistical dependence should be valid without any assumptions of an underlying probability density function and should be extensible to high dimensionality of input measurements. A recent approach combines an adaptive Gaussian sum filter with the Kullback-Leibler (KL) divergence measure for effective data association [1]. However, the KL divergence does not satisfy all the properties of a true distance metric, making analysis of results all the more challenging, in addition to the fact that computing the KL divergence is computationally demanding. In general, we want the chosen statistical dependence metric between two observations to identify a nonlinear higher-than-second order dependence between measurements, in order to claim a pair-wise association between observations. In our previous work, we demonstrated that mutual information is the most promising metric of statistical dependence for OTOA, which confirms other findings in the literature [2], [3].

In this paper, we revisit a simple OTOA problem with two closely spaced RSOs. We first describe the overall procedure and how it relates to initial orbit determination (IOD). Second,

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

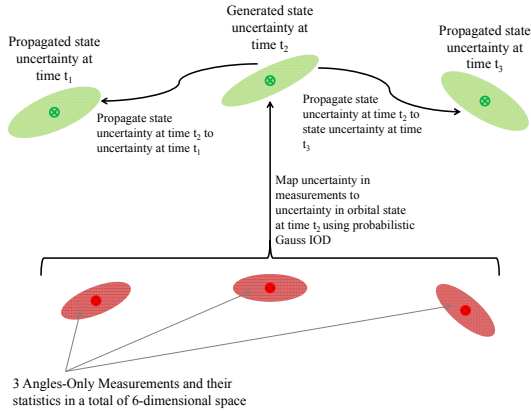


Fig. 1. The general probabilistic IOD approach developed by Hussein et al. [4].

we describe the derivation and calculation of the IOD correction. Third, we demonstrate the application of the correction in LEO, MEO, GTO, and GEO orbit regimes and its benefit to the mutual information data association approach. Finally, we summarize the main results of the paper.

II. IOD AND THE OTOA PROBLEM

The core idea in the proposed OTOA approach is to use an appropriate IOD method to generate an uncertain track from a set of measurements and their known statistics that we wish to test for association (see Fig. 1). One can then compare, in some information theoretic sense, the amount of information shared between the generated orbit statistics and the measured output statistics. The more consistent the estimated track is with the measurements, the more likely the observations were generated by the same physical phenomenon. This method of comparison is based on the notion of mutual information between the IOD-based orbit statistics and the measured observation statistics (see Fig. 2). In previous work [2], it was found that mutual information performed consistently better than other information theoretic metrics. Therefore, we focus our attention exclusively on mutual information in this paper.

For illustration purposes, we will assume that the observations are angles-only pairs of azimuth and elevation. We will use the classical Gauss method for the IOD step (see [5]). In general, we assume we are given a set of n observations $\{z_1, \dots, z_n\}$ taken at observation times t_1, \dots, t_n (times are assumed to be distinct, without any loss of generality). Of these, we will choose three observations (as required by the Gauss method) to test whether or not they were generated by the same RSO. The procedure is applied to all combinations of three observations. How to computationally address the combinatorics problem is beyond the scope of this paper, and we will only address the problem of determining whether a set of three observations were generated by the same RSO. In the general n observation problem, those observations that

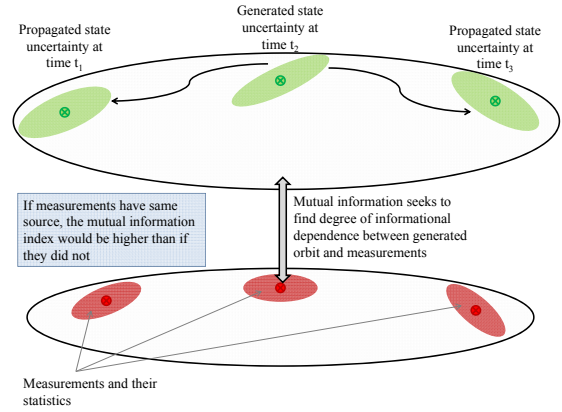


Fig. 2. Mutual information can be used as an index of how much dependence exist between a set of measurements and the orbit they generate if they were associated. The “more” unassociated the observations are the smaller the mutual information will be, and vice versa.

were deemed to be associated would next form tracks, and the remainder of this problem becomes a TTTA problem that proceeds by “stringing” associated tracks together to form a set of non-redundant new tracks. We do not address the TTTA problem in this paper.

Focusing on the three measurements $z = (z_1, z_2, z_3)$ taken at t_1, t_2 , and t_3 , the Gauss IOD method produces a candidate orbit described by the six-dimensional state $x_2 = g(z)$ at time t_2 , where $g(\cdot)$ is the function that maps a set of three angles-only measurements to orbital space coordinates. The state may be specified in orbital elements, Cartesian coordinates, etc. Furthermore, let f_{ij} be the function that propagates the state x_i defined at time t_i to the state x_j defined at time t_j . Then $x_1 = f_{21}(x_2)$ is the (backward) propagated state at time t_1 and $x_3 = f_{23}(x_2)$ is the propagated state at time t_3 given the state x_2 at time t_2 . Finally, let h_i be the function that maps the state at time t_i to an observation $z_i = h_i(x_i)$.

III. IOD CORRECTION METHODOLOGY

Based on the results of our previous work, the performance of the proposed data association schemes depends on the performance of the underlying IOD method. Therefore, it stands to reason that if we can improve the IOD methodology, then we would improve our ability to use mutual information for data association. In previous work, Wilkins [6], [7] explored corrections to atmospheric density models and found that it was possible to use TLEs as pseudo-observations of the atmosphere to back out corrections to the density model. With a similar thought in mind, we explore the idea that the IOD methodology can be corrected using angles-only observations as pseudo-observations of the IOD process.

In the case of dynamic calibration of the atmosphere, the error in the density modeling process showed up as structured noise in estimates of the ballistic factor. One could take

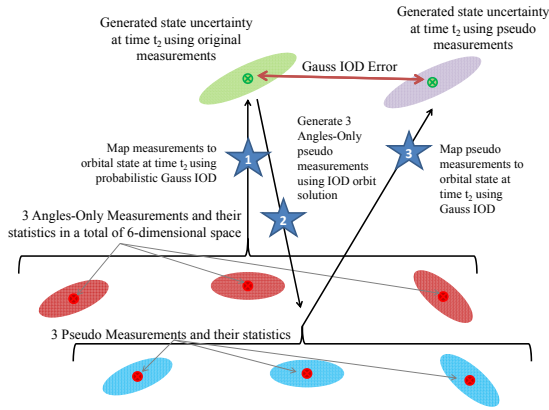


Fig. 3. The probabilistic IOD correction approach

estimates of the ballistic factor and compare to some “truth” ballistic factor to generate a correction as a function of a variety of parameters. For our purposes, there is no directly observable or estimable quantity by which we can measure the outcome of the IOD process. Furthermore, we will never know the “true” orbit for comparing our initial IOD solution. What we can do is construct a process that will allow us to create an approximation of the error in the IOD process. This process is depicted in Fig. 3.

We assume that the unknown true orbit at t_2 is given by \mathbf{x}_2^t . To begin, use the angles measurements $\mathbf{z} = (z_1, z_2, z_3)$ to generate an IOD solution for the orbital state at time t_2 as per usual. Let this solution be denoted by \mathbf{x}_2^e at time t_2 . Our intuition is that, while the IOD solution is erroneous, it is a deterministic process that will produce a specific error for a given input. We wish to calculate an IOD correction by determining an estimate of $\mathbf{x}_2^e - \mathbf{x}_2^t$. To that end, we need a second IOD orbit solution from a set of observations generated by a true orbital state in the neighborhood of \mathbf{x}_2^e , which we denote by \mathbf{x}_2^e . We contend that the difference between the two IOD solutions $\mathbf{x}_2^e - \mathbf{x}_2^e$ is representative of the error introduced by the IOD process for $\mathbf{x}_2^e - \mathbf{x}_2^t$ itself. To construct \mathbf{x}_2^e , we use the IOD solution \mathbf{x}_2^e to generate a new set of pseudo angles measurements $\mathbf{z}^e = (z_1^e, z_2^e, z_3^e)$ and our existing sensor models.

There are more sophisticated approaches for computing corrections along the lines of the dynamic calibration of the atmosphere research. However, for our initial purpose of determining a simple bias correction, the current method is sufficient to illustrate the benefits of such an approach. Further research is warranted into how this IOD correction is generated.

IV. ANALYZING ORBIT UNCERTAINTY

In order to analyze the uncertainty in the orbital space resulting from a given uncertainty in the measurement space, we will use the unscented transform (UT) according to the

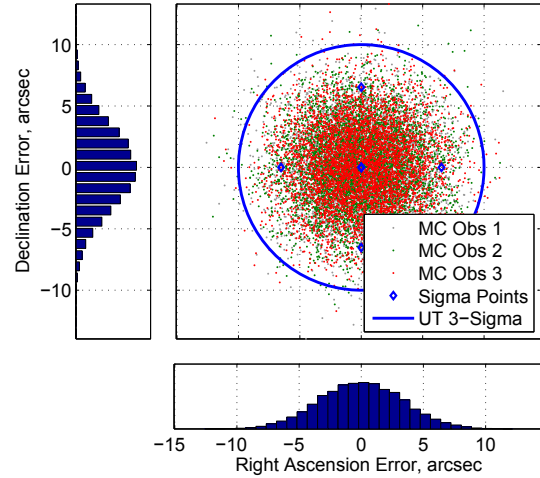


Fig. 4. Example of sampling a six-dimensional Gaussian measurement uncertainty.

procedure described by Hussein et al. [4]. For this UT analysis, the measurement process is assumed to be Gaussian and the generated orbit uncertainty will then be modeled as a Gaussian. There are a total of $13 = 2 \times 6 + 1$ sigma points $\{\mathcal{Z}^{(j)}\}$ since the measurement space has a dimension of 6. The resulting Gaussian distribution in the orbital space is an approximation of the actual uncertainty pdf $p_O^2(\mathbf{x}_2)$ resulting from the IOD solution.

As noted by Hussein et al. [4], there is a subtle difference between this description of the problem compared to other approaches which have appeared previously in the literature. The nonlinear mapping $\mathbf{g}(\cdot)$ is a map from the entire three-measurement space to the orbital space, and is *not* a map of an individual measurement z_i , $i = 1, 2, 3$, to the orbital space. Therefore, samples of the measurement uncertainty should be drawn from the distribution in the six-dimensional measurement space defined with the global measurement variable \mathbf{z} and not from the individual distributions defined on the individual measurement variables z_i , $i = 1, 2, 3$. For the UT method, in particular, this will result in the correct number of sigma points (13) being generated to describe the uncertainty distribution in six-dimensional orbital space.

For example, a Monte Carlo sample and UT sigma points are shown in Fig. 4 for a Gaussian uncertainty with each dimension independently distributed with $3\sigma = 10$ arcsec. Each colored set of dots represents a two dimensional projection of the MC sample onto each individual measurement plane (i.e., there are $5000 \times 3 = 15\,000$ dots shown, but the MC sample only contains 5000 particles). The remaining dimensions of the sigma points overlap each other since the uncertainty is the same in each direction. The histograms show the marginal distributions of the MC azimuth and elevation uncertainty samples for one of the measurement times.

V. MUTUAL INFORMATION CRITERION FOR OTOA

The mutual information $I(\mathbf{x}, \mathbf{z})$ between two random variables \mathbf{x} and \mathbf{z} is a measure of the degree of dependence between these two variables. Formally, it is given by

$$I(\mathbf{x}, \mathbf{z}) = \iint p(\mathbf{x}, \mathbf{z}) \log \left(\frac{p(\mathbf{x}, \mathbf{z})}{p(\mathbf{x})p(\mathbf{z})} \right) d\mathbf{x}d\mathbf{z}, \quad (1)$$

where $p(\mathbf{x}, \mathbf{z})$ is the joint distribution of \mathbf{x} and \mathbf{z} , and $p(\mathbf{x})$, respectively $p(\mathbf{z})$, is the marginalization of $p(\mathbf{x}, \mathbf{z})$ with respect to \mathbf{z} , respectively \mathbf{x} . Two important properties are to be noted here. Firstly, mutual information is a symmetric function of \mathbf{x} and \mathbf{z} . Secondly, if \mathbf{x} and \mathbf{z} are independent then $p(\mathbf{x}, \mathbf{z}) = p(\mathbf{x})p(\mathbf{z})$ and the mutual information is zero (immediately follows from (1)).

The mutual information index we propose to use is defined as follows. First, we consider the mutual information $I_i(\mathbf{x}_i, \mathbf{z}_i)$ between the state \mathbf{x}_i and the measured output \mathbf{z}_i at time t_i . The overall mutual information index would then be $I_{\text{tot}} = I_1 + I_2 + I_3$. Other indices based on mutual information could also be considered. For example, one can consider the mutual information between the *joint* orbital state variable $\mathbf{x}_{\text{joint}} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ and the joint measurement variable $\mathbf{z}_{\text{joint}} = (\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)$. Such indices will be the subject of future research.

It can be shown that the mutual information $I_i(\mathbf{x}_i, \mathbf{z}_i)$ can be expressed in terms of the KL divergence:

$$I_i(\mathbf{x}_i, \mathbf{z}_i) = D_{KL}(p(\mathbf{x}_i, \mathbf{z}_i) || p(\mathbf{x}_i)p(\mathbf{z}_i)) \quad (2)$$

where D_{KL} is the KL divergence between two pdfs $p(\mathbf{x})$ and $q(\mathbf{x})$, which is given by

$$D_{KL}(p(\mathbf{x}) || q(\mathbf{x})) = \int p(\mathbf{x}) \log \frac{p(\mathbf{x})}{q(\mathbf{x})} d\mathbf{x} \quad (3)$$

If both p and q are Gaussian, then one can compute $D_{KL}(p || q)$ in closed form (see, for example, [1]):

$$D_{KL}(p(\mathbf{x}) || q(\mathbf{x})) = \frac{1}{2} \left[\log \left(\frac{\|\Sigma_q\|}{\|\Sigma_p\|} \right) + \text{Tr}(\Sigma_q^{-1} \Sigma_p) - d + (\boldsymbol{\mu}_p - \boldsymbol{\mu}_q) \cdot \Sigma_q^{-1} \cdot (\boldsymbol{\mu}_p - \boldsymbol{\mu}_q) \right], \quad (4)$$

where $\boldsymbol{\mu}_p$ and Σ_p (resp., $\boldsymbol{\mu}_q$ and Σ_q) are the mean and covariance of the pdf p (resp., q), and d is the dimension of the underlying space.

Following the UT procedure described in [4], one can obtain a Gaussian approximation of the pdf of the state \mathbf{x}_2 . One can then use the UKF to obtain the joint distribution $p(\mathbf{x}_i, \mathbf{z}_i)$ after propagating and updating with the corresponding measurement \mathbf{z}_i , $i = 1, 2, 3$. It can be shown that this distribution is Gaussian with the following mean:

$$\boldsymbol{\mu}_i^{\text{joint}} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{x}_i} \\ \boldsymbol{\mu}_{\mathbf{z}_i} \end{bmatrix}$$

and covariance:

$$\Sigma_i^{\text{joint}} = \begin{bmatrix} \Sigma_{\mathbf{x}_i} & \Sigma_{\mathbf{x}_i, \mathbf{z}_i} \\ \Sigma_{\mathbf{x}_i, \mathbf{z}_i}^T & \Sigma_{\mathbf{z}_i} \end{bmatrix},$$

where $\boldsymbol{\mu}_{\mathbf{x}_i}$ is the UKF *updated* state mean at time t_i , $\boldsymbol{\mu}_{\mathbf{z}_i}$ is the measurement mean at time t_i , $\Sigma_{\mathbf{x}_i}$ is the *updated* covariance in the state at time t_i , $\Sigma_{\mathbf{z}_i}$ is the measurement covariance at time t_i , and $\Sigma_{\mathbf{x}_i, \mathbf{z}_i}$ is the *updated* cross-covariance between the state and the measurement at time t_i . The reason we use the *updated* statistics as opposed to the *propagated* statistics is that we are, in the first place, assuming that the measurements are associated (the hypothesis to be tested).

Since the joint pdf is Gaussian, the marginalization with respect to \mathbf{x}_i and \mathbf{z}_i are both Gaussian and have means $\boldsymbol{\mu}_{\mathbf{x}_i}$ and $\boldsymbol{\mu}_{\mathbf{z}_i}$, respectively, and covariances $\Sigma_{\mathbf{x}_i}$ and $\Sigma_{\mathbf{z}_i}$, respectively. Now that we have Gaussian approximations of $p(\mathbf{x}_i, \mathbf{z}_i)$ and $p(\mathbf{x}_i)p(\mathbf{z}_i)$, we can use (2) and (3) to compute the mutual information $I_i(\mathbf{x}_i, \mathbf{z}_i)$.

VI. SIMULATION RESULTS

In this simulation, we compare the performance of the OTOA solution, with and without corrections, in four orbit regimes: LEO, MEO, GTO, and GEO. For testing, we consider two objects, with identification numbers 0 and 1, in close proximity to one other. They both have identical orbital elements, listed in Table I, except for the value of the true anomaly at the initial time. One object has the value listed in the table while the other is modified by a small delta. A set of angles-only observations of the two RSOs are collected at three different times. If we arbitrarily index the two measurements with 0 and 1, then the question is which sequence of tags are the correct ones? There are eight possible combinations of tags: 000 (i.e., observations with tags 0 at the three time instances are associated to one of the two objects and so on), 001, 010, 100, 011, 101, 110 and 111. The observations were tagged such that the two correct ones are 000 (all coming from RSO number 0) and 111 (all coming from RSO number 1). The mutual information criterion used in our previous work [2] is used here again: how close can the two objects be in true anomaly before the method being tested fails to return the two correct associations as the two most likely ones? Note that when a method “fails,” while the correct associations may not be the one most highly ranked, they would rank very close to the top. As the separation distance between the two objects decreases further, the correct associations are farther from being top ranked, and they are more or less arbitrarily ranked as all solutions become indistinguishable.

The measurement model has three basic parameters: sensor location, time between observations, and angular measurement noise. The sensor latitude and longitude used for each case are provided in Table I. The topocentric azimuth and elevation observations were collected at the rate of one observation every N minutes with the value of N listed in the table (also notice the large measurement error standard deviation chosen for these simulations). The measurement noise is assumed to be Gaussian with an angular standard deviation for both azimuth and elevation specified in Table I.

For each of the cases, the separation in true anomaly between the two RSOs is steadily decreased until the method fails to report the correct associations as the most likely

TABLE I
PARAMETERS OF THE TRUE ORBIT AND MEASUREMENT MODEL

Parameter	LEO	MEO	GTO	GEO
Semimajor Axis, km	6 991.	26 600.	24 461.	42 157.
Eccentricity	0.0	0.2	0.7322	0.01
Inclination, deg	97.9	55.	19.3	0.0
Argument of Perigee, deg	0.0	-120.	-90.	0.0
Right Ascension of the Ascending Node, deg	-62.8	-106.7	-17.5	0.0
True Anomaly, deg	-30.	110.	200.	260.8
Measurement Noise σ , deg	0.67	2.	10.	3.33
Time Between Observations, min	2.5	20.	20.	60.
Sensor Latitude, deg	20.7088	33.8172	-7.41173	60.7088
Sensor Longitude, deg	-156.2578	-106.6599	72.45222	-156.2578

TABLE II
TRUE ANOMALY DIFFERENCE AT WHICH ASSOCIATION METHOD FAILS

Orbit Regime	IOD Corrected	True Anomaly Difference (deg)
LEO	No	5.96108
LEO	Yes	0.20465
MEO	No	0.30997
MEO	Yes	0.01546
GTO	No	1.02704
GTO	Yes	0.78725
GEO	No	1.40198
GEO	Yes	0.27058

ones. The association picked by each method is the one that produces the maximum value of mutual information index. As can be seen in Table II, the correction we introduced to the mutual information OTOA solution results in an order of magnitude improvement in association performance for the cases considered. It is important to note that the performance results will depend on changes in many variables, including time between measurements and orbital geometry. The performance improvement was greatest for the LEO case, followed by the MEO and finally by the GEO and GTO cases, though the MEO case was the one with the best performance, possibly due to the chosen set of parameters. This is not surprising given the degradation of the Gauss IOD method as one approaches the GEO belt. It is important to note that while the resolution capability of this OTOA method is relatively large (~ 54 arcsec for the MEO case) compared to typical sensor resolutions, one has to remember that the OTOA problem being solved is far more challenging than the classical OTTA problem. With the method proposed in this paper, and others to be considered in future work, we can sift through a collection of observations, not correlated to any object or uncorrelated track, and sort out which ones are most likely to belong together given minimal knowledge of the observation statistics.

VII. CONCLUSION

In our prior work, it was shown that the mutual information method performed the best among the information theoretic approaches that were tested for the observation-to-observation association (OTOA) problem. In this paper, we tested a variation of the mutual information technique, with the addition of a correction factor for the IOD method. Simulation results indicated an order of magnitude improvement in the overall OTOA problem performance using this correction over the uncorrected method. Future work will examine whether we need to capture IOD error as a function of other parameters. In the long run, it is generally conjectured that Monte Carlo-based methods will prove to be more efficient as they more faithfully represent the underlying probability distributions. This paper barely touches the tip of the iceberg, and its most basic goal is to induce the space community to further consider the use of information theoretic measures for solving the various data association problems (OTOA, OTTA, and TTTA).

REFERENCES

- [1] R. Linares, V. Kumar, P. Singla, and J. L. Crassidis, "Information theoretic space object data association methods using an adaptive gaussian sum filter," *Advances in the Astronautical Sciences*, vol. 140, pp. 665–680, 2011, (Proceedings of the 21st AAS/AIAA Space Flight Mechanics Meeting, New Orleans, LA, February 13–17 2011, Paper AAS 11-148).
- [2] I. I. Hussein, C. W. T. Roscoe, M. P. Wilkins, and P. W. Schumacher, Jr., "Information theoretic criteria for observation-to-observation association," in *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference*, Wailea, HI, September 9–12 2014.
- [3] A. Alempijevic, S. Kodagoda, and G. Dissanayake, "Mutual information based data association," in *Proceedings of the 5th IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, Melbourne, Australia, December 7–10 2009, pp. 97–102.
- [4] I. I. Hussein, C. W. T. Roscoe, M. P. Wilkins, and P. W. Schumacher, Jr., "Probabilistic admissibility in angles-only initial orbit determination," in *Proceedings of the 24th International Symposium on Space Flight Dynamics*, Laurel, MD, May 5–9 2014.
- [5] P. R. Escobal, *Methods of Orbit Determination*. John Wiley & Sons, Inc., 1965, (Reprint with corrections by Krieger Publishing Company, 1976).
- [6] M. P. Wilkins, C. Sabol, P. J. Cefola, and K. T. Alfriend, "Validation and application of corrections to the NRLMSISE-00 atmospheric density model," *Advances in the Astronautical Sciences*, vol. 127, 2007, (Proceedings of the 17th AAS/AIAA Space Flight Mechanics Meeting, Sedona, AZ, January 28–February 1 2007, Paper AAS 07-189).

- [7] —, “Discourse on corrections to the NRLMSISE-00 atmospheric density model,” in *Proceedings of the 7th US-Russian Space Surveillance Workshop*, Monterey, CA, November 2007.