Multi-Phenomenology Fusion for Satellite Identification

Trevor Putman
Senior Professional Staff, JHU/APL

Major Caleb McComas, Lieutenant Colonel Julia Faustman
Space Security Defense Program

Abstract

The space domain is becoming increasingly crowded with proliferated constellations and advanced, highly-maneuverable satellites. New mission sets, such as on-orbit servicing, introduce space scenarios that stress our current Space Domain Awareness (SDA) approach. These complex space scenarios include large maneuvers, rendezvous proximity operations (RPO), and deployment events. Maintaining accurate identification of satellites is a critical component of the SDA mission. Traditionally, space operators identify satellites by associating their kinematic information with a known catalog. The trend of increasing complexity in the space domain suggests a kinematic approach may prove insufficient to maintain identification of space objects in the near future. There is a compelling need to provide space operators with the ability to leverage information beyond traditional kinematics when determining the identity of orbital objects.

As the complexity of space operations increases, so does the quantity and availability of SDA data. Commercial SDA systems enable lower-latency, high-frequency collection of kinematics and unresolved signature information on satellites in all orbital regimes. Unresolved signature information refers to sensor-measured information such as radar cross section, visual magnitude, and passive radiofrequency collections. While individual sensors may use non-resolved signature data to distinguish objects within their field of view, complex space scenarios indicate a need to incorporate multiple phenomenologies to achieve an accurate characterization.

This study explores the use of unresolved signature information across multiple phenomenologies to identify satellites with a calculated confidence. The result of this study will be a proposed framework and classification algorithm for using multiple phenomenologies to classify a satellite. Additionally, the study highlights future force design and data exposure recommendations to improve SDA systems' capabilities for classifying satellites. The study defines an algorithmic and architectural approach to providing a classification capability as a decision aid to SDA Command and Control (C2) system operators.

This study explores multiple variations of Naïve Bayesian classifiers for use in satellite identification. Bayesian classifiers are particularly well suited to this problem set due to their ease of explanation, low computational complexity, and statistics-based confidence values. They are also robust to incomplete datasets which is a critical need in the data-sparse space domain environment. We will train and evaluate the classification algorithms against real-world scenarios using historical datasets pulled from the Unified Data Library (UDL). Our algorithm's objective is to label the orbital object with the correct satellite catalog number based on a comparison of current observed signature information with historical-based signature profiles.

Our results indicate significant utility in using unresolved signature information to identify satellites. Bayesian classifiers prove to be a simple yet effective method to leverage this information, providing a 95% accuracy across historical scenarios. Furthermore, analysis of the reported confidence identifies clear thresholds for using the output of the algorithm as a decision aid to space operators. A sensitivity analysis of contributing phenomenologies highlights the utility of each phenomenology for providing a classification call. Furthermore, the sensitivity analysis highlights key areas for investment to improve satellite identification capabilities.

In this study, we establish the usefulness of unresolved signature information, but space operators need to be able to convert these signatures into actionable decision aids. Our study defines a clear path forward to provide operators with defendable identification calls in complex scenarios.

Introduction

The space domain serves a critical role in communication, position, navigation & timing (PNT), environmental monitoring, and many other mission areas. Space is a keystone of global economy and an area of significant investment in the commercial, civil, and government sectors [1]. Over the past decade, this domain has evolved to host record numbers of satellites [6]. The U.S. Department of Defense (DoD) describes the space domain as "increasingly congested, contested, and competitive [2]". The introduction of large proliferated Low Earth Orbit (pLEO) constellations, such as Starlink and OneWeb, further increases the congestion of space. New emerging mission sets such as on-orbit servicing have the potential to increase spacecraft lifetimes and enable complex space missions [3]. On-orbit refueling missions require satellites to enter extremely close proximity, often referred to as Rendezvous Proximity Operations (RPOs). These scenarios can complicate efforts to track and distinguish satellites.

New mission areas introduce complex scenarios for agencies tasked with maintaining Space Domain Awareness (SDA). The United States Space Force (USSF) defines SDA as "encompassing the effective identification, characterization, and understanding of any factor associated with the space domain that could affect space operations. [4]" Maintaining accurate identification of satellites is a critical component of the SDA mission and a primary focus of this paper. The evolution of the space domain to a congested, dynamic environment will make accurate identification more challenging and more imperative.

The process of satellite identification falls into two main methodologies: cooperative and non-cooperative. The cooperative methodology leverages cooperation with satellite owner/operators to maintain position and identification of a space craft via satellite communication and Telemetry, Tracking and Command (TT&C). The non-cooperative methodology uses external sensors to characterize a space object and compare it to an established space catalog. This paper focuses on non-cooperative identification of orbital objects.

Traditionally, non-cooperative space object identification has been done using exclusively kinematics. The success of this approach hinges on two key assumptions: orbital objects are spaced far apart, and orbital objects do not perform large maneuvers to change their orbit. Current trends in space operations challenge these assumptions. For example, the Chinese satellite Shijan-21 docked with a defunct Beidou satellite and towed it into super-synchronous orbit roughly 300 km above the geosynchronous (GEO) belt [7]. This scenario highlights how future on-orbit servicing operations will involve orbital behavior (close approaches, docking, and large maneuvers) that challenges long-held SDA assumptions.

In tandem with the growing complexity of the SDA mission, commercial companies are investing heavily in their own SDA capability. Commercial companies operate geographically dispersed sensors across multiple phenomenologies to provide SDA data to government and commercial stakeholders. These sensors include, but are not limited to, ground-based telescopes, ground-based radar, and passive radiofrequency receivers. Commercial SDA systems enable lower-latency, high-frequency collection of kinematics and unresolved signature information on satellites across all orbital regimes. Unresolved signature information refers to sensor measured information such as radar cross section, visual magnitude, and passive radiofrequency collections. While individual sensors may use non-resolved signature data to distinguish objects within their field of view, complex space scenarios indicate a need to incorporate multiple phenomenologies to achieve an accurate characterization.

Currently, much of this commercial SDA data is aggregated in the Unified Data Library (UDL). The USSF Chief of Space Operations (CSO) designated the UDL as the data library for storing and accessing this commercial data in support of USSF needs [5]. Furthermore, the United States Government Accountability Office (GAO) recommends the USSF leverage commercial data in support of its SDA mission [5]. This paper explores how fusion of unresolved, commercial signature data can support a critical component of the SDA mission - space object identification.

Approach

In this section, we will discuss our algorithmic approach enabling space object identification.

Data Availability

When considering how to address the challenge of identifying objects in space, we need to fully examine the information available to support an identification decision. The following is a list of the information leveraged in this paper:

Kinematics: Information detailing the location and velocity of a target space object in time.

Radar Cross Section (RCS): The apparent size of an object derived from the reflected power collected at a radar. This will vary depending on the aspect of the target relative to the measuring sensor and the frequency of the sensor.

Visual Magnitude (Vmag): The measured intensity of an object detected by a passive optic. Vmag is represented in the log-scale. This will vary depending on the aspect of the target relative to the measuring sensor. These measurements are fit to a diffuse spherical model and normalized in both range and solar phase angle.

Passive Radiofrequency: The frequency of signals emitted from a target space object. These signals are collected by passive radiofrequency detectors for the purpose of Time/Frequency Difference of Arrival (TDOA/FDOA) tracking.

Polarity: The measured polarization of signals emitted from a target space object. This information is collected by determining peak power received in a particular polarization: horizontal, vertical, right-hand circular, left-hand circular.

The above data types are a subset of a much broader range of possible information. The study selected these features due to their widespread availability in commercial datasets. More advanced products such as resolved imagery from optics and inverse synthetic aperture radar could be extremely beneficial for object identification but are currently out of scope.

Algorithm Selection

While the space domain is experiencing significant growth in both quantity and complexity of orbital operations, man-made orbiting objects have existed for decades. The long orbital lifetime of spacecraft allows for sensor collections across large temporal and geographic baselines. Aggregated historical signature information for an individual spacecraft provides a good understanding of expected measured signatures in the future. We are presented with a classification problem where large historical signature databases exist for the objects we wish to classify. These signature databases can be refined into measured distributions known as priors. Figure 1 shows an example of a discrete distribution of measured signals.

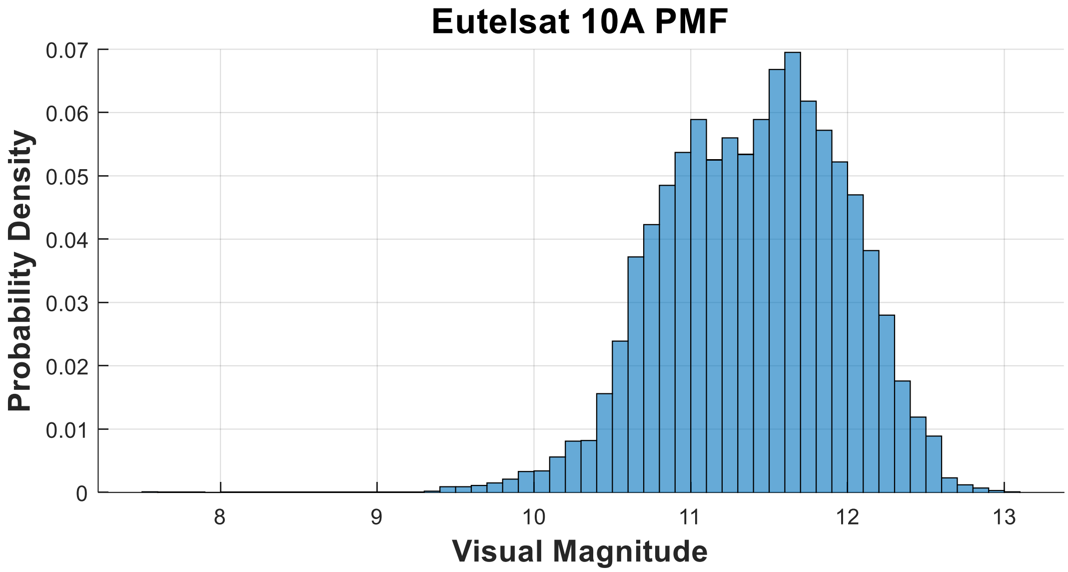

Figure 1 VisMag distribution for Eutelsat 10A

Bayesian classifiers are a family of classifiers built on Bayes Theorem, and they are well suited to this form of classification problem. Theses algorithms compare measured data points against prior distributions to determine which distribution most closely matches the data points. Equation 1 highlights the Bayes Theorem applied to a satellite identification problem where a ground-based optic has measured a Vmag of 11.

$$P(A \mid B) = \frac{P(B \mid A) * P(A)}{P(B)} \quad (1)$$

Where,

P(A|B) is the probability a measured satellite is Eutelsat 10A given a measured Vmag of 11

P(B|A) is the probability of measuring a Vmag 11 on Eutelsat 10A

P(A) is the probability of the measured object being Eutelsat 10A

P(B) is probability of the sensor measuring a Vmag 11

The Bayesian classifier discriminates objects by comparing their relative likelihoods. This approach assumes no bias in the measuring sensor such that P(B) is a uniform value for all sensor measurements. Additionally, we assume a uniform value for P(A) such that any satellite (meeting our kinematic gating threshold) is equally likely to be the measured satellite. We are left with a comparison of P(B|A) to determine the relative likelihood that a measured object is any particular satellite. We derive P(B|A) by referencing the Probability Distribution Function (PDF) of the prior distribution. We accumulate the likelihood of any particular satellite being the measured object by taking the product of P(B|A) over each measurement (as seen in Equation 2).

$$Likelihood_{Eutelsat10A} = \prod_{i=1}^{n} P(B_i \mid A) \quad (2)$$

Where,

n is the number of measurements

$P(B_i \mid A)$ is the probability that measurement B_i was measured from Eutelsat 10A

Taking the product of probabilities results in a value rapidly trending towards zero. To address this challenge, we store the likelihood values as negative logarithmic likelihoods. When the algorithm provides an output, it will normalize the likelihoods among the candidate satellites such that they will sum to one. Normalization provides a much more intuitive confidence value than extremely small negative values.

We chose Bayesian classifiers for several reasons. Most notably, these algorithms are extremely robust to disparate data sets. They can produce a classification decision from a single data point and any combination of phenomenology. This is especially important when ground-based sensors have limited access to certain spacecraft over long periods of time. Additionally, Bayesian classifiers are computationally lightweight and explainable. A consumer of these classification decisions could compare prior distributions with current measurements to understand the statistical decision the algorithm made.

Representing the Prior

In our approach, we construct a prior distribution per orbital object for each phenomenology. Unfortunately, these distributions do not typically fall within the known and easily modeled distributions (ex. Gaussian). Numerous factors (including aspect angles, odd satellite configurations, specular glint, etc.) create irregularities in these prior distributions. We propose three methods to account for the irregularity of these prior distributions. The first method involves storing the distribution as a normalized, discrete probability distribution function as seen in Figure 1. Discretizing the prior is computationally simple and allows the algorithm to only store limited information (bins and counts). However, the selection of bin width and placement is completely subjective. Furthermore, grouping datasets in bins can smooth key discriminant behavior in an observed prior. For these reasons, method 1 is considered sub optimal to method 2.

Method 2 models the prior by fitting a Gaussian Mixture Model (GMM) to the distribution. GMMs allow us to model a complex distribution as a sum of Gaussian components. Similarly to method 1, this method only requires the algorithm to store limited information (means, sigmas, and mixing proportions) for each satellite prior. However, method 2 allows for much greater resolution in the actual distribution of the prior and preserves key discriminant behavior. Figure 2 shows the probability distribution function of a GMM fit to a prior.

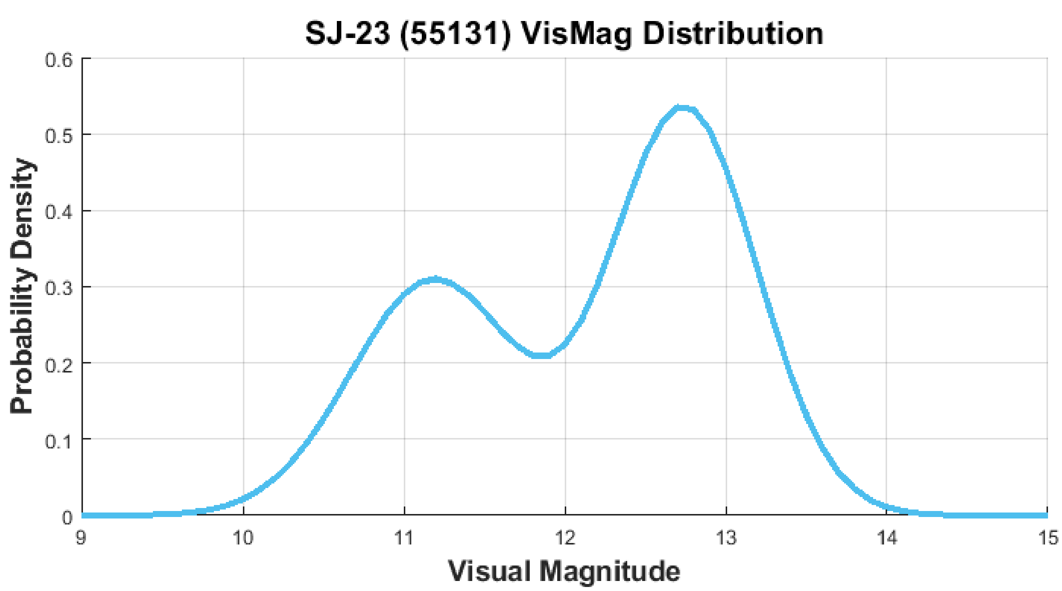

Figure 2 SJ-23 Vmag distribution represent by Gaussian Mixture Model

While GMMs are well suited to modeling continuous distributions of data, they are particularly ill suited to modeling categorical data (ex. polarity). Therefore, we propose a third method, which is a hybrid of method 1 and 2. Method 3 models priors of continuous features as GMMs and priors of categorical features as discrete PDFs.

All three methods will be implemented and explored in the following sections.

Algorithm Functional Flow

Figure 3 shows the functional flow of the classification algorithm. Note that sensor data, state estimation, and the space catalog are all considered outside the scope of this identification algorithm. State estimation is the process of correlating sensor measurements and performing an orbit determination. Signature data must be correlated to an orbital object to be fused for object identification. For our purposes, correlation should only consider kinematic information (ex. Right Ascension, Declination for optics). If signature data is used in the correlation process (feature-aided correlation), our object identification approach becomes incestuous. Signature data cannot be used twice; once to correlate a measurement to a specific satellite and the second time to identify the satellite. Double counting the signature information results in a self-confirming bias.

We designed the identification algorithm to be hosted inside a broader Command & Control (C2) algorithm that would correlate measurements to orbital objects and produce states. The referenced catalog could be produced by the hosting C2 algorithm or could leverage existing catalogs such as the one on SpaceTrack.org. The following subsections will cover the sub-components of the identification algorithm.

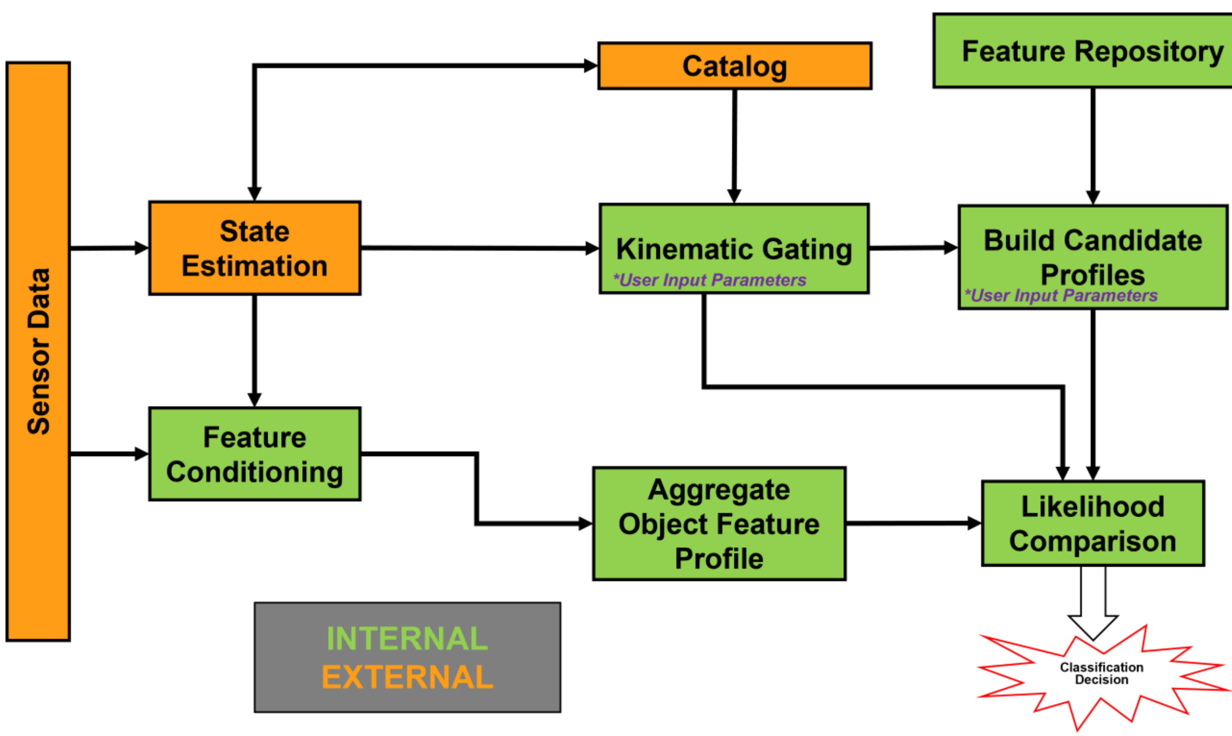

Figure 3: Functional flow diagram for identification algorithm

Kinematic Gating

The first step in our classifier scopes down the list of possible candidates using kinematics. The kinematic gating component determines which orbital objects could have maneuvered to be the unknown object. This function outputs the optimal velocity expenditure (delta-V) needed for each known object to maneuver to the unknown object state. This is additional information a consumer could leverage when making a classification decision. This function outputs a list of orbital candidates that fall under some user-set velocity threshold.

Feature Conditioning

Concurrently to the kinematic gating step, the algorithm normalizes all incoming signature information for use in the downstream classifier. RCS values are normalized based upon the appropriate region (optical, Rayleigh, Mie). Vmag measurements are normalized in both range and solar phase angle using a diffuse sphere phase function. Current features pulled from passive radiofrequency do not require any normalization.

Aggregate Object Feature Profile

This step stores all of the collected feature information prior to feeding the data to the Bayesian classifier. Storing this data allows the algorithm too update the feature repository after the classification is complete.

Build Candidate Profiles

The algorithm pulls all available priors for each candidate and phenomenology present. Priors are updated every day to incorporate new signatures into the modeled distribution. The priors are stored in a feature repository.

Likelihood Comparison

This step is the application of the Bayesian classifier. It compares the aggregated signature information currently being measured from the target with the historical prior distributions. This step outputs a classification call and an associated confidence/likelihood for that call.

Results

To find scenarios for test, we employed a separate algorithm to identify maneuvers and RPO events based on historical Two-Line Element sets (TLE). The algorithm replayed these scenarios using the sensor correlation (satellite number tag given by sensor) and the catalog on space-track.org as inputs. All presented results used real-world data on orbital objects. We tested this algorithm against 666 real-world scenarios including maneuvers and RPO events. Truth data was the classification call made by Space-Track.org when a new TLE was published on an object post-maneuver or post-RPO.

The algorithm ingested scenario data sequentially as it would if it was running in a live environment. It produced a classification call and associated confidence each time a measurement was ingested. All three methods for representing the prior were tested and the results are shown in Table 1.

Algorithm	% of Objects Correctly Identified
Discrete Naïve Bayesian Classifier	**90% (597 / 666)**
GMM Bayesian Classifier	**95% (632 / 666)**
Hybrid Bayesian Classifier	**95% (632 / 666)**

Table 1: Scenario results

The value in the second column represents the number of scenarios in which the algorithm arrived at the correct classification decision and did not change its classification decision with additional information. The GMM and Hybrid approaches (Methods 2 and 3) demonstrated the best performance. Given the Hybrid classifier's flexibility to better handle categorical features, the Hybrid classifier appears to be the best solution.

Figure 4 displays the algorithm output confidence compared with the percentage of algorithm calls that were correct in truth. Classification confidence is computed by normalizing the calculated likelihood for each candidate satellite. Normalized likelihood values are much more intuitive for end users than raw negative log likelihood values.

Ideally, the data would display a linear pattern where the reported confidence accurately represented the likelihood the classification call was correct. However, our results showed variations in the correlation between reported confidence and correctness. This behavior was particularly noted in the middle confidence bins. When the algorithm reports a confidence higher than 90% it is correct roughly 99% of the time. With this behavior in mind, consumers of the classification call could use 90% as suitable threshold for trusting the output of the classifier. Ultimately, the output confidences are rooted in statistics and forensic analysis can explain how the algorithm arrived at both its decision and confidence.

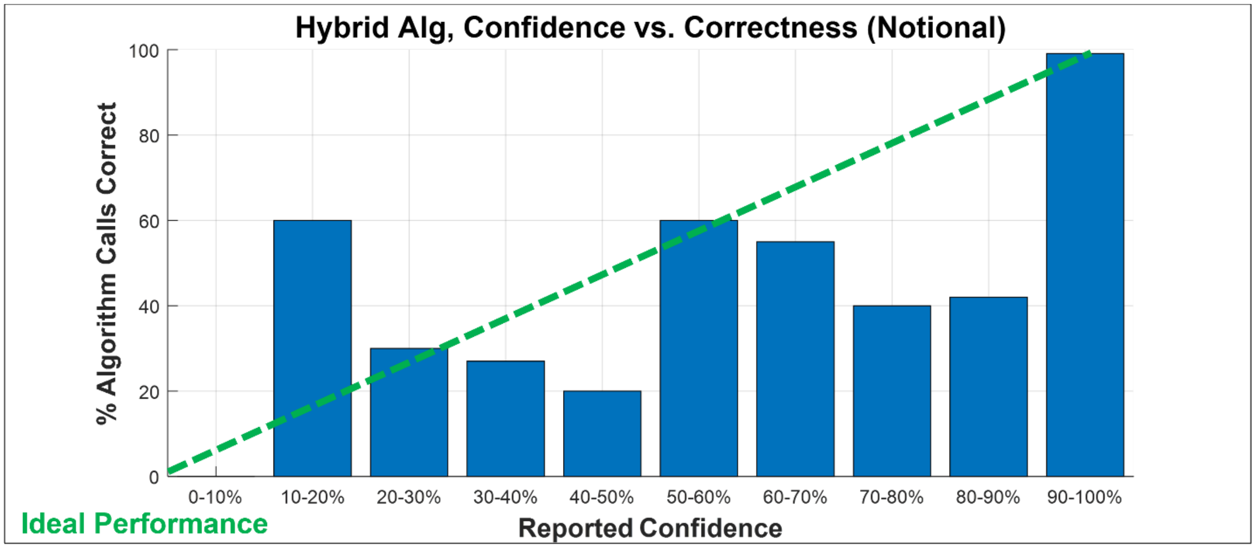

Figure 4: Classification confidence versus correctness

Analysis of the cases in which the classifier arrived at the incorrect answer identified data sparsity as a key limitation. Apart from a single scenario, all of the cases of incorrect classifications had very little data for the classifier to ingest. Figure 5 shows a breakdown of the data availability for the incorrect classification cases.

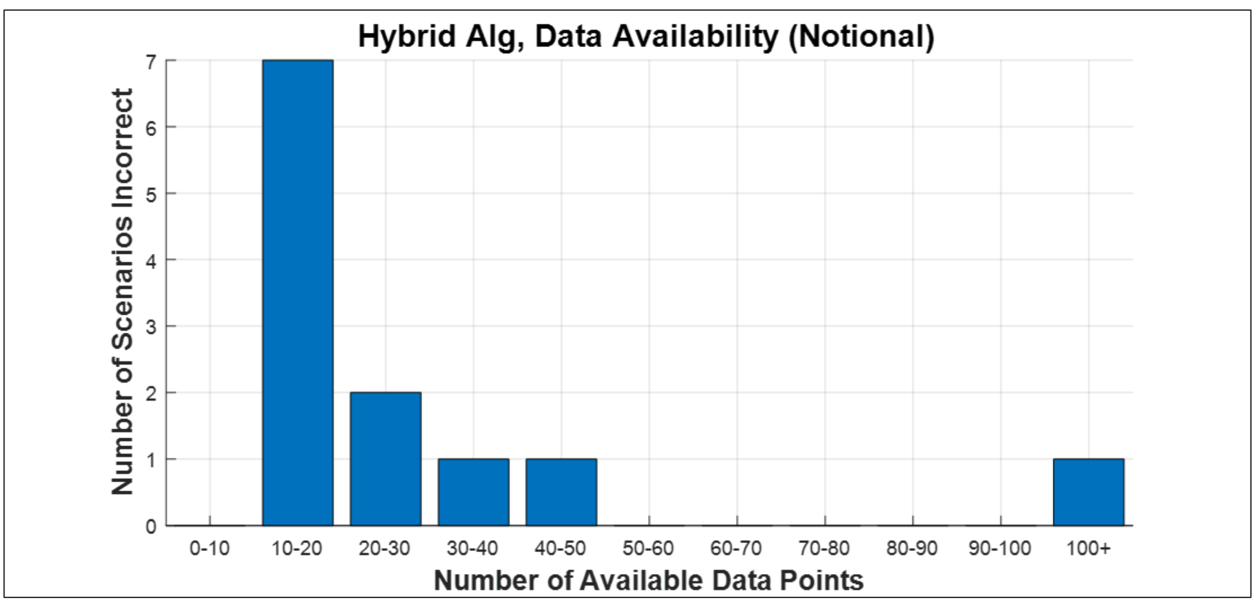

Figure 5: Data availability for incorrect cases

In addition to evaluating the algorithm's performance against real-world data, we performed a sensitivity analysis. The sensitivity analysis removed individual phenomenologies from each scenario and regraded the algorithm performances. This analysis highlights which phenomenologies are most impactful for object identification. Table 2 shows the changes in algorithm performance when certain phenomenologies are removed.

Phenomenology Removed	% of Scenarios with Performance Degradation	% of Scenarios with Performance Improvement
PASSIVE Frequency	80%	2%
Radar Cross Section	75%	7.4%
Polarity	60%	1.5%
Visual Magnitude	28%	5.1%

Table 2: Sensitivity analysis results

In general, removal of any phenomenology results in a net degradation of algorithm performance. This behavior reinforces the need for fusion across phenomenologies to provide accurate identification. Passive radiofrequency proved to be the most useful phenomenology with radar cross section following closely. This analysis highlights opportunities to invest in passive radiofrequency and radars to improve enterprise object identification capabilities.

There were limited scenarios where removal of a phenomenology improved the algorithm performance. These are cases where objects appeared similar in feature space for that phenomenology, and inclusion of that phenomenology injected uncertainty into the algorithm.

Overall, the hybrid algorithm performed very well when identifying orbital objects in real-world scenarios. The output confidence values displayed a reasonable threshold to use when trusting classification calls, and the algorithm settled on the correct answer in 95% of the scenarios.

Additional Applications

Our approach of modeling historical signature distributions using GMMs enables many other applications. A space C2 algorithm could compare the GMMs of multiple space objects to determine the overlap/correlation between expected distributions. Fig. 6a shows the GMM Vmag distributions for eight satellites. Fig. 6b quantifies the overlap of these distributions. Satellite pairs with a high value have very similar distributions, and pairs with a low value are easily separable in feature space.

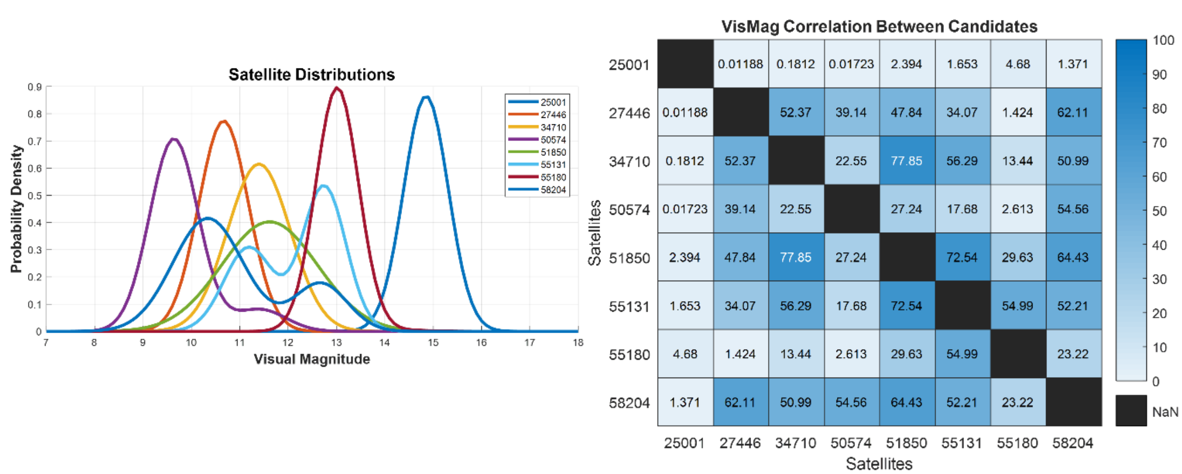

Figure 6: (a) Left-side, GMM VisMag distributions for eight satellites

(b) Right-side, Correlation between VisMag Distributions

For example, satellite 25001 (represented by the rightmost blue curve in Fig 6a) has a much higher Vmag distribution than the other satellites. By referencing the first column of Fig 6b, we can see low correlation values indicating Vmag is a good phenomenology to identify 25001 among these satellite options. This approach could be paired with a sensor orchestrator to prioritize tasking of phenomenologies that are most impactful in distinguishing closely spaced objects. Signature-informed tasking would improve enterprise SDA performance and free up sensor resources to support missions for which they are well suited.

GMM-based signature distributions also support anomaly identification. An anomaly identification algorithm compares incoming signature data, correlated with a known orbital object, to an established historical distribution. By comparing this data with a GMM representation, we can derive the statistical distance of a measurement from its known distribution. Consumers of anomaly identification alerts could set statistical thresholds for alerts (ex. two-sigma) and leverage statistical distances to quantify the degree of the anomaly.

Finally, feature-aided correlation is a topic of interest in the SDA community. Feature-aided correlation includes feature information in the cost function used by a correlation algorithm. Our GMM approach creates numerical likelihood values for each measured signature and a corresponding orbital object. A correlator would use these values in the cost function to weight correlation of a measurement towards the object it more closely matches in signature space. Feature-aided correlation is a powerful tool for minimizing cross-tagging of measurements and maintaining accurate tracks of closely spaced objects. As previously noted, this identification algorithm should not be paired with feature-aided correlation because it creates a self-confirming bias.

Limitations

Algorithm performance is affected by several variables including scene complexity, available phenomenologies, and data volume. The algorithm struggles to distinguish similar objects in close proximity. In these cases, passive radiofrequency data is an ideal discriminant to identify objects, but this data is not always available.

Additionally, the Bayesian classifiers perform well if the current measured signature information matches historically measured signatures. Spacecraft that change their physical configuration may confuse the classifier. In cases where historical signatures have not been captured for a known space object, the algorithm represents their distribution as a normal distribution centered around the global median value. This approach can lead to sub-optimal classifications when objects identified as candidates have not been historically characterized.

Recommendations

Our study demonstrates the utility of applying Bayesian classifier to the space object identification challenge. We hypothesized the need for data fusion to support this mission, and the sensitivity analysis provided confirming evidence. Data exposure and aggregation are key elements supporting the space object identification mission. Data repositories, such as the Unified Data Library, are key enablers to accessing commercial data across various phenomenologies. Due to our algorithm’s low computational complexity, the timeliness of data aggregation drives the timeliness of identification call. Entities interested in space object identification should prioritize the timely exposure of data to support identification timelines.

While our test scenarios are based on real-world data, sensors were not deliberately tasked to collect on our objects of interest. Data used in the scenarios was collected by commercial sensors following their own tasking schedule. Pairing this identification algorithm with a sensor orchestrator would improve timeliness, data availability, and collection of specific, beneficial phenomenologies. Additionally, our approach is based on establishing an accurate historical model of the range of possible signatures measured from an object. A sensor orchestrator paired with our algorithm could enable prioritization of sensor collections from various aspect angles to build a robust historical dataset. Object identification and characterization are critical elements of the SDA mission and a global sensor orchestrator should integrate those capabilities into its tasking priorities.

Conclusion

An "increasingly congested, contested, and competitive [2]" space environment requires novel approaches to address the challenge of identifying objects in orbit. Bayesian classifiers demonstrate significant utility in classifying orbital objects. These classifiers provide an explainable identification call and a statistics-based confidence value. The proposed hybrid Bayesian classifier achieves 99% accuracy when its reported confidence is above 90%. Furthermore, the GMMs representing the historical signature databases show promising utility for applications in anomaly detection, feature-aided correlation, and signature-informed tasking. Sensitivity analysis of the Bayesian classifier indicates investments in passive radiofrequency and radar sensors could improve enterprise object identification efforts.

In this study, we established the usefulness of unresolved signature information for object identification, but space operators need to be able to convert these signatures into actionable decision aids. Our approach defines a clear path forward to provide operators with defendable identification calls in complex, space scenarios.

References

[1] Bland, Rob & Brukardt, Ryan & Gangware, Will & Swartz, Dale. 16 November 2022. *A different space race: Raising capital and accelerating growth*. Retrieved from: https://www.mckinsey.com/industries/aerospace-and-defense/our-insights/a-different-space-race-raising-capital-and-accelerating-growth-in-space

[2] Department of Defense. 2011 January. *National Security Space Strategy*. Retrieved from: https://dod.defense.gov/Portals/1/features/defenseReviews/NSSS/NationalSecuritySpaceStrategyUnclassifiedSummary_Jan2011.pdf

[3] Davis, Joshua & Mayberry, John & Penn, Jay. 2019 April. *On-Orbit Servicing: Inspection, Repair, Refuel, Upgrade, and Assembly of Satellites in Space*. Retrieved from: https://aerospace.org/sites/default/files/2019-05/Davis-Mayberry-Penn_OOS_04242019.pdf

[4] Headquarters United States Space Force. 2022 January. *Operations Doctrine for Space Forces*. Retrieved from: https://media.defense.gov/2022/Feb/02/2002931717/-1/-1/0/SDN%20OPERATIONS%2025%20JANUARY%202022.PDF

[5] U.S. Government Accountability Office. 2023 April 24. *Space Situational Awareness: DOD Should Evaluate How It Can Use Commercial Data*. Retrieved from: https://www.gao.gov/products/gao-23-105565

[6] McDowell, Jonathan. 2023 Jan 17. *Space Activities in 2022.* Retrieved from: https://planet4589.org/space/papers/space22.pdf

[7] Jones, Andrew. 2022 January 27. *China's Shijian-21 towed dead satellite to a high graveyard orbit.* Retrieved from: https://spacenews.com/chinas-shijian-21-spacecraft-docked-with-and-towed-a-dead-satellite/