Autonomous, hybrid space system fault and anomaly detection, diagnosis, root cause determination, and recovery

Dick Stottler, Sowmya Ramachandran, Chris Healy, Abhimanyu Singhal,
Stottler Henke Associates, Inc.

ABSTRACT

An important component of Space Situational Awareness (SSA) / Space Domain Awareness (SDA) is knowledge of the true status of friendly assets and whether any assets are under attack. Therefore, it is important to be able to detect faults and other anomalies, and determine the components involved and the root cause as well as whether that root cause is likely an external attack. During space conflict, communications to satellites may be disrupted, requiring them to intelligently and autonomously "take care of themselves," i.e., effectively detect faults, diagnose their root causes, and develop and execute recovery plans, autonomously, without necessarily being able to communicate with ground controllers. This lack of communication is analogous to lunar rovers and power systems where communication can be disrupted by terrain and other factors.

Astrobotic, for NASA, is developing a rover that traverses over the lunar surface to an advantageous position, then unfurls a 60' high photovoltaic mast to provide power for other lunar systems. Astrobotic's Vertical Solar Array Technology (VSAT) will egress from its lander, transit to the desired location (near the lunar South Pole), "wiggle" into the lunar soil, and then deploy a 60' high solar array to generate and then distribute power to other lunar systems. The VSAT will include several subsystems, such as mobility, internal and external (to provide power to external systems) electrical power systems, thermal management, and array deployment, each of which must work smoothly in order for the operation to succeed. As the VSAT moves around the surface of the Moon, sensors are constantly providing information on how much traction is available and how quickly the rover is moving. As the solar array is unfurled, a gimbal system and inertial measurement units (IMUs) continuously monitor the array's movement, including any lean. If the array leans too much, the solar array can buckle—worse, the entire rover may be at risk of tipping over, failing the mission. Since the array is so tall compared to VSAT's wheelbase, even just a few degrees of lean would be disastrous. This situation may be very dynamic, denying ground controllers enough time to correct any problem, given the round trip communication delays.

It is therefore important that the VSAT be equipped with the means to quickly detect problems, perform diagnosis and root cause determination, and quickly safe the system. Traditionally, Fault Detection, Isolation, and Recovery (FDIR) systems have utilized Model Based Reasoning (MBR), which requires knowledge of the subsystem design and the behavior of components down to the desired level of diagnosis. To the degree this information is readily available, it is important to make good use of it. However, the field of machine learning (ML) has shown that systems can also learn, offline, the normal behavior of complex systems in many different environments and states, and then detect abnormal behavior in real time. These systems can also be trained with known abnormal states, and recognize these more specifically when they occur.

This paper will describe progress on this work since our last paper, presented at AMOS 2022. This includes further development and generalization of the hybrid approach to fault detection, diagnosis, and recovery as well as how we are applying that approach to the most critical aspects of VSAT.

With the new types of subsystems (such as mechanical components and related sensors) came new challenges to be overcome. Some concerns included quick reaction times needed to avoid tipping or buckling during mast deployment and, at the opposite end of the spectrum, detecting very gradual changes, hard to discern in sensor noise (the mast moves very, very slowly while tracking the Sun). In some cases, data is severely limited, reducing the applicability of a pure ML approach.

These challenges led to the development of a third, independent method for detecting anomalies, based on an analogy to thermodynamic variables: the Thermodynamic Reasoning for Intelligent Anomaly Detection (TRIAD) system, which performs automatic Characterization and Diagnosis of subsystem anomalies. Similarly to how actual thermodynamic values such as pressure and temperature help to summarize in macro form the condition of a large number of micro aspects (e.g., the speeds of individual molecules), TRIAD loosely uses this same concept to

summarize groups of sensor values. Examples include mean or variance over the last N received datapoints for a sensor. Others are min, max, and max jump between two samples. Additional functions utilize Fourier Transforms (FTs) of the incoming data stream, generating added variables such as average frequency, min frequency, max frequency, peak frequency, and amplitude of peak frequency. Note that every generating function is defined over the last N samples so that one function may be "mean over the last 100 samples" and another may be "mean over the last 500 samples." Because looking at multiple time scales can be helpful, TRIAD maintains multiple "versions" of each type of function where each version corresponds to a different N.

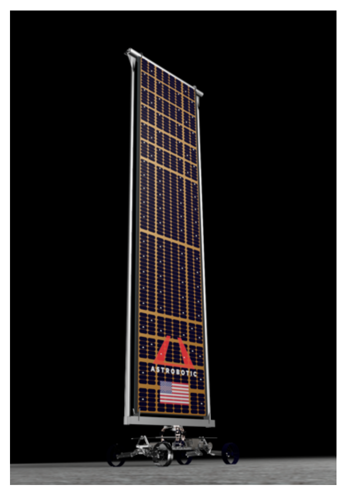
The hybridization emphasizes the benefits of each approach and mitigates the disadvantages. The benefits of the hybrid system include the ability to detect and diagnose anomalies never before encountered; working well on "Day One" of operations; effectively utilizing existing design knowledge; succeeding without large amounts of data; explaining the reasoning and being human understandable; being rigorously certifiable; behaving predictably; diagnosing down to the lowest modelled component level; handling rare but modeled operating conditions; executing very quickly; discovering unknown and subtle relationships (even across subsystems); and finally, providing extra certainty of the diagnosis when all three approaches agree.

1. MOTIVATION

As the U.S. Government and private companies continue to develop and launch spacecraft, there is a growing need to increase the level of autonomy of each system, especially in an adversarial world. One of the key capabilities for an autonomous system is the ability to manage faults and off-nominal behavior. If a mechanical failure occurs in space, it is usually impossible for a human to be able to fix the issue; systems that can manage faults autonomously free up time and effort from ground controllers, enabling human beings to focus on potentially more pressing issues. Systems may also be outside of line of sight (LOS) at the time of the fault, due to being in eclipse relative to human operators or terrain preventing direct communication.

A general fault management (FM) architecture is needed to combat faults across a range of spacecraft subsystems. Management in a fault scenario for a given subsystem would entail first detecting that a problem has occurred, then safing the spacecraft to minimize any damage, diagnosing the problem, and determining the root cause; it next calls for determining what courses of action (COAs) might be feasible given the current situation (failed sensors, mechanical failures, etc.) and selecting the best one, then generating a schedule or sequence of actions to implement the COA, and finally, executing those actions. However, FM is a vigilant process and must be proactive in monitoring system health.

Rovers on the Moon are examples of semi-autonomous systems that will need to be equipped with sophisticated and robust FM systems. One such rover is Astrobotic's Vertical Solar Array Technology (VSAT), which will egress from its lander, transit to the desired location (near the lunar South Pole), "wiggle" into the lunar soil, and then deploy a 60' high solar array to generate and then distribute power to other lunar systems. A graphic of the VSAT is shown on the right. The VSAT will include several internal subsystems, such as propulsion, internal electrical system, thermal management, and array deployment, each of which must work smoothly in order for the operation to succeed. As the solar array is unfurled, a gimbal system, inertial measurement units (IMUs), wheel load sensors, inclinometers, and up-facing camera continuously monitor the array's movement, including any lean. If the array leans too much, the array could buckle—worse, the entire rover may be at risk of tipping over and failing the mission. Since the array is so tall compared to VSAT's wheelbase, even just a few degrees of lean would be disastrous.



2. BACKGROUND

As previously reported, we have applied these techniques to a number of spacecraft subsystems, including the EPS of the Power Propulsion Element (PPE) and Habitation and Logistics Outpost (HALO) modules of the Lunar Gateway (in simulation), a fault-tolerant processing and EPS monitoring experiment onboard the ISS, the CO2 Removal Loop and Thermal Management System on the External Portable Life Support System (xPLSS) (in simulation), Johnson Space Center's (JSC's) simulation of the Mars Transit Vehicle (MTV) [8], Montana State University's (MSU's) LabSat hardware (a more complex version of their in-space CubeSat IT-SPINS), the ISS's Urine Processing Assembly (UPA), and the NASA Ames Graywater Recycling System (GRS) [7]; and integrated them with NASA's core Flight System (cFS).

We are working with Astrobotic, the developer for VSAT, to develop, for ultimate fielding on the Moon, a system we call Modular AI for Faults: Local Online Watch and Efficient Response (MAIFLOWER). VSAT is envisioned to be deployed on a lunar crater near the southern pole, where line of sight communications may be severely limited (making VSAT a perfect target for MAIFLOWER fault detection).

MAIFLOWER additionally employs a model-free anomaly detection and diagnosis module, TRIAD, as an extension of previous efforts. TRIAD is a modular system that serves as a potent anomaly detector on its own but can additionally incorporate other time-series anomaly detection frameworks. TRIAD focuses on the *intelligent aggregation* of time-series anomaly detection methods and can detect and diagnose faults that no individual method can analyze on its own; as such, as the field advances, TRIAD's modular design enables the efficient incorporation of new state-of-the-art methods to maximum effect. While we have previous used machine learning methods for fault detection, MAIFLOWER (via improvements to TRIAD) is the first to explore time series anomaly detection using transformers. The present state-of-the-art methods for time-series anomaly detection utilize deep-learning architectures such as transformers and Convolutional Neural Networks (CNNs) in order to transform time-series data into easily actionable diagnoses.

3. SYSTEM OVERVIEW

Here we describe MAIFLOWER at a high level, followed by a more detailed description in Part 4. During normal operations, MAIFLOWER monitors onboard sensor values to automatically characterize subsystem components and to be prepared to detect failures. During a failure scenario, MAIFLOWER would first detect the problem; immediately safe the spacecraft to minimize damage; then diagnose the problem and determine the root cause.

Any intelligent, adaptive system must inherently be a closed loop system (i.e., in basic terms: the system must sense what is occurring and make appropriate decisions to take suitable actions, and sense the effects of those actions). The first part of this sense-decide-act loop involves perception, understanding the situation from the raw sensor values. Over long-time scales, this involves characterization. The other perceptual function is detecting faults and diagnosing their causes.

Model-based reasoning (MBR) systems are often used as in these types of applications; these systems encode the schematic information of subsystems, which includes the components (including sensors), their normal behavior and known abnormal modes of behavior, and the connections between components. During normal operations, the model is used to simulate the current behavior and compare the simulated sensor output values to the actual sensor outputs. Significant deviations are used to detect some kind of fault. Then the model is used to reason which component faults are most likely to lead to the currently deviating sensor values. The set of possible faults (possibly including sensor faults) which explain the sensor values is the MBR diagnosis engine's output. The process of using the model to diagnose failures is considered somewhat analogous to the reasoning an engineer would employ when using a schematic to try to diagnose the fault. The process can be made more efficient by various heuristics used by spacecraft engineers to quickly diagnose problems and include knowledge of which components are most likely to fail (and how) and/or are the most likely explanation for certain types of sensor values. MAIFLOWER takes advantage of these heuristics.

In addition to MBR, MAIFLOWER leverages an expansion of the TRIAD system to synthesize an extensive model-free module to detect and diagnose faults. TRIAD has shown excellent performance in fault detection and classification across fault types with regard to several spacecraft subsystems. Beyond this, the TRIAD system constitutes an overarching framework that can intelligently incorporate and synthesize any set of additional feature-based anomaly detection systems. Feature-based anomaly detection systems utilize deterministic functions to extract time series of low dimensional features from incoming streams of sensor data and perform simple threshold-based anomaly detection on the extracted series. Nearly every prominent anomaly detection algorithm can be classified as feature-based anomaly detection, from state-of-the-art deep learning algorithms using transformer and Convolutional Architectures [1][2][5][10] to more traditional algorithms making use of Hidden Markov Models or Self-Organizing Maps [3][4]. TRIAD is designed to intelligently synthesize these algorithms instead of running them in parallel.

The advantages of TRIAD's synthesis are twofold. First, TRIAD's synthesis is significantly more accurate than any parallel implementation, for purposes of both detection and characterization. Second, TRIAD enables more wholistic modelling of the utility of each method's individual contribution to the accuracy of the system. For example, a computationally slow/expensive anomaly detection method may outcompete cheaper methods, run individually or in parallel, but find itself redundant when incorporated into TRIAD's system. By contrast, a cheaper method may fail to provide any novel detection when compared against already-implemented methods but significantly enhance expressivity when incorporated in intelligent concert with these methods within TRIAD.

MAIFLOWER extends TRIAD with feature functions from cutting-edge time-series anomaly detection algorithms, including the state-of-the-art methods that utilize transformer and CNN architectures. Transformer and CNN architectures have broadly dominated the state-of-the-art for anomaly detection in recent years. Deep learning encoders are trained on a variety of downstream tasks, including detection of synthetic faults [5], the prediction of future data [2], and reconstruction accuracy [1]. The feature encodings are then subject to threshold-based detection. We are investigating a "mix-and-match" method of encoding architectures (transformers, CNNs) and downstream tasks, incorporating each into TRIAD in order to comprehensively assess and implement the most robust and efficient combinations of state-of-the-art techniques.

4. MAIFLOWER SYSTEM DESCRIPTION

Shown in the figure to the right is the high-level architecture for MAIFLOWER. In general, MAIFLOWER can support fully autonomous, closed loop execution. The loop is constantly monitoring sensor values across subsystems to ensure all systems are running nominally. However, when sensor values appear to significantly deviate from typical conditions, MAIFLOWER first uses information from its various sensors to determine the likelihood of a fault occurring (as opposed to a false positive). MAIFLOWER exploits various techniques (model based reasoning, ML, and thermodynamic analysis, among others) to evaluate the probability of a fault and attempt to diagnose the situation. Once a fault has been pinpointed, MAIFLOWER quickly safes the system. In later versions, it will also move on to replanning goals, such as the need to reconfigure or take other actions to restore mechanical systems or power after a fault and execute those operations. However, our initial version of MAIFLOWER is specifically focusing on Characterization, Anomaly Detection, Diagnosis, and Safing.

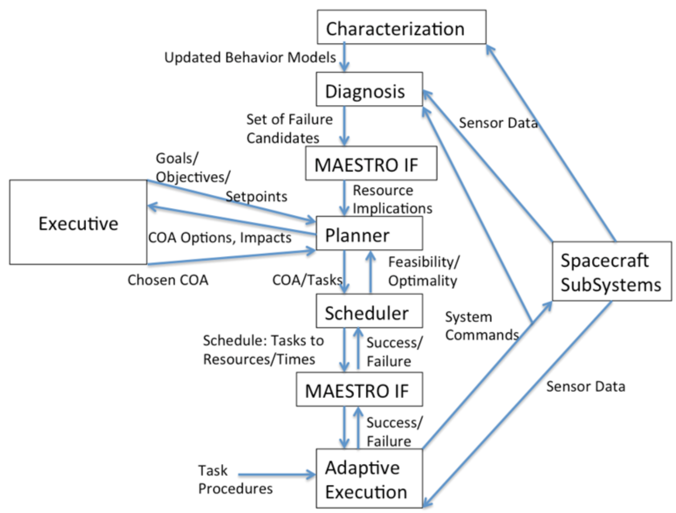


Figure 1: MAIFLOWER High-Level Architecture.

5. MODEL-BASED REASONING (MBR) DIAGNOSIS ENGINES

Model-based diagnosis systems encode detailed and explicit descriptions of the interrelated factors that affect phenomena. These models typically represent the world as a collection of components, where each component is characterized by attribute values and one or more modes. Constraints specify required relationships among attribute values and modes, and constraint violations are used to identify components in faulty modes. For example, forces, torques, and translational and angular accelerations, velocities, and positions, and different sensor values of these quantities, are constrained by algebraic relationships and equations of motion. If the sensed values do not obey these relationships, either one of the sensors or the physical component must be at fault.

MBR systems have traditionally been used to diagnose faults in engineered systems. System components are characterized by nominal and known faulty modes, and constraints on interconnected components are based on physical laws. Because models encode the effects of contextual factors, they can be applied reliably across contexts, such as the current environment, configuration, and sent commands. Model-based reasoning requires knowledge engineering efforts to encode these interacting effects. MBR engines can be extremely fast and do not require a large amount of memory or compute power, even for complex models.

During normal operations, the model is used to simulate the current behavior and compare the simulated sensor output values to the actual sensor outputs. Significant deviations are used to detect some kind of fault. The model is then used to reason which component faults are most likely to lead to the deviating sensor values. The set of possible faults (including sensor faults) which explain the sensor values comprise the MBR diagnosis engine's output. As touched on earlier, the process of using the model to diagnose failures is considered somewhat analogous to the reasoning an engineer uses when seeking to diagnose a fault by use of a schematic. The process can be made more efficient by various heuristics which spacecraft engineers use to quickly diagnose problems and include knowledge of which components are most likely to fail and how, and/or are the most likely explanation for certain types of sensor values. MBR engines identify one specific fault and/or a set of possible faults. Note that the MBR engine does this every time new data arrives. However, as mentioned later, it may wait to issue a detection until other points in time or other modules have been consulted. Being hosted onboard the VSAT allows several sets of telemetry data to be received each second and processed, which can be used to minimize false alarms due to sensor noise, even within the 1-second reaction requirement.

6. THERMODYNAMIC REASONING FOR INTELLIGENT ANOMALY DETECTION (TRIAD)

The central idea behind TRIAD is that of a Thermodynamic Variable (TD). Note that in this context, a TD is a name for a mathematical object and does not necessarily refer to quantities such as temperature and pressure that are associated with thermodynamics specifically. This concept is inspired by thermodynamics but not limited to it. A TD is a function of the most recent n datapoints, with n varying for each TD. A simple example of a TD is the variance across the last 100 received datapoints. TRIAD recursively applies sets of generating functions to input data in order to generate maximally expressive sets of TDs. Different sequences of functions can be handcrafted applications or developed via optimized sampling. Examples of functions used in the previously include mean, min, max, variance, and max jump between two samples. Additional functions utilized Fourier Transforms (FTs) of the incoming data stream, generating additional TDs such as average frequency, min frequency, max frequency, peak frequency, and amplitude of peak frequency. Note that every generating function is defined over the last n samples. Therefore, one function may be “mean over the last 100 samples” and another may be “mean over the last 500 samples.” Because looking at multiple time scales can be helpful, TRIAD maintains multiple “versions” of each type of function where each version corresponds to a different n.

TRIAD leverages fault-free training data to construct quantile-based empiric distributions over each of the generated TDs. When the system is online, out-of-distribution instances of features are detected as anomalies. The thresholds for what constitutes an out-of-distribution value for variables can be tweaked depending on the desired sensitivity, moving along a tradeoff frontier between false positive and false negative rates. One can further extend this tweaking to adjust the sensitivity to specific *types* of faults so as to remain hypervigilant for catastrophic faults while reducing the false alarm rate for less dramatic events.

To characterize anomalies, TRIAD leverages training data of past or simulated anomalies to develop vectorized recordings of feature deviations for each anomaly event in the training data. TRIAD can then fit probability distributions in this space for each fault type by maximizing the log-probability of parameterized gaussian mixture distributions. Gaussian mixtures are highly expressive distributions, especially in contexts where target distributions can be high-dimensional and multi-modal. This allows for incomplete sets of “cues” for a specific type of fault to remain actionable, which can be helpful in the case of broken sensors. Using these optimized distributions, TRIAD can assign a probability to each class of known anomaly in the event of a real fault, in addition to a probability that the fault is of an unknown type. Both the unsupervised anomaly-detection component and the supervised anomaly-characterization component can be retrained with new data in the face of new operational circumstances with no development overhead, allowing TRIAD continually learn and keep pace with new contexts, incorporating new knowledge of specific fault types over time.

To incorporate additional anomaly detection methods *into* TRIAD, one just inputs the associated feature-generating function into the pool of TRIAD’s generating functions. TRIAD starts at baseline with an implementation identical to parallel usage of each of its incorporated anomaly detection methods; TRIAD monitors the outputted time-series generated by each incorporated feature-generating function as if they were TDs, essentially replicating the model. However, TRIAD can then improve upon this method with the sequential application of feature-generating functions. For example, an incorporated CNN network that does not shed light on the raw input data may consistently detect otherwise-overlooked anomalies when run on the time-series of Fourier features extracted by TRIAD. Alternatively, the Fourier features of the time series data extracted by a transformer architecture may alert the system to anomalies that the transformer would otherwise overlook. These are examples of sequences of two generating functions: TRIAD supports manually designed sequences and can be extended to optimize and curate function sequences of arbitrary length to minimize the redundancy of monitored values up to the computation budget allotted. As such, for any given computation budget, TRIAD has the capability to provide a far more comprehensive monitoring system than a baseline parallel implementation.

7. MACHINE LEARNING FOR ANOMALY DETECTION

Deep learning architectures such as transformers and CNN have broadly dominated the state-of-the-art for anomaly detection in recent years. Transformer architectures have revolutionized Natural Language Processing (NLP) tasks in the last several years but have yielded improved performance over other networks on nearly every task involving time-series data, such as financial price-history datasets or video footage. CNNs have been the standard architecture for Computer Vision for many years and continue to broadly dominate the state-of-the-art in that field. CNNs have also been adapted to one-dimensional data such as time-series to great effect—while individual CNN layers are less expressive than the attention layers used in transformers, they are computationally faster, and as such, CNNs can be significantly deeper at the same computational cost.

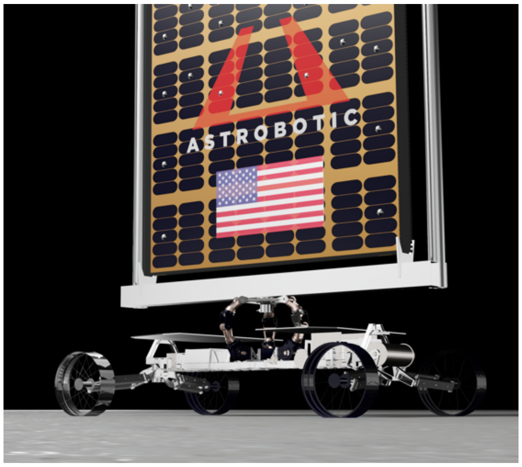
The use of trained networks of time-series anomaly detection is a matter of standard classification when sufficient labelled data of anomalies exists in abundance. When anomalous data is scarce, the networks must be instead trained on alternate downstream tasks adjacent to the task of anomaly detection, such that the time-series of encodings generated for that task are conducive to simpler anomaly detection methods. Many downstream tasks have been proposed, and the selection of downstream tasks constitutes the central innovation of most state-of-the-art literature on this topic. Prominent downstream tasks in the current frontier of performance are the detection of future datapoints in the stream [2] and autoencoding [1]. Models trained for autoencoding (called autoencoders) are trained to generate a lower-dimensional representation of input features, while a decoder is simultaneously trained to recreate the input features from the lower-dimensional representation. Both networks are trained to maximize the accuracy of the decoder. Finally, many algorithms generate synthetic fault data and train standard classification algorithms. While deep learning architectures have outcompeted the prior state-of-the-art for anomaly detection by a wide margin, no clear downstream task has emerged as the general standard for training, and new downstream tasks are introduced frequently.

8. HYBRIDIZING MBR AND TRIAD

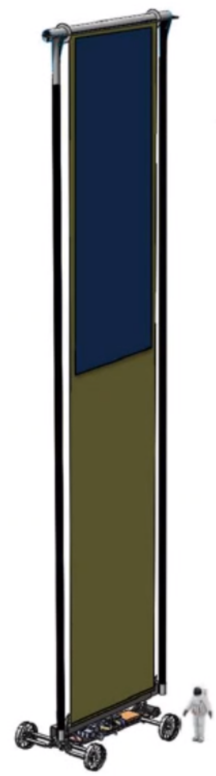
A hybrid fault detection system uses both TRIAD and MBR to monitor sensor values in an effort to identify anomalous behavior. Hybridizing MBR and TRIAD emphasizes the benefits of each while minimizing their disadvantages. Generally, the benefits of MBR coincide with the disadvantages of TRIAD, and vice versa, so the two complement each other well [3]. When an anomaly is detected, a few possibilities present themselves: both systems detect the anomaly, only TRIAD detects it, or it is only detected by the MBR system. TRIAD detects all data streams that are "different" from the data on which it was trained, and as such it generates reports when encountering novel outside contexts, even when all systems are functioning effectively. MBR is a good check on these false alarms, especially if a method exists to determine if TRIAD training data does not include data from the current operating state. However, TRIAD's model-free anomaly detection can also detect faults that escape the analytic limits of MBR. Because MBR and TRIAD are very different technologies (especially in that one is model-based while the other is model-free), when they agree, this consensus instills extra confidence in the result (compared to using MBR or TRIAD alone).

9. ASTROBOTIC'S VSAT SYSTEM

By way of example—as NASA intends to dramatically increase the number of manmade objects and systems on the Moon, it is essential to develop a grid to generate and distribute electrical power. This grid will consist of several elements, from the generators to the intermediate distributors to the final recipients, and a multitude of faults could occur at any stage in this process. Astrobotic's Vertical Solar Array Technology (VSAT) is an example of a generator platform. At a high level, it resembles a rover with a large vertical solar array that can unfurl once the rover has reached a suitable area for solar power generation and that rotates to track the motion of the Sun. The VSAT is likely to be deployed near the lunar south pole, which is why it is oriented vertically. A graphic of the array is shown lower right, while the figure below provides a close-up view of the rover.



The VSAT platform will face several challenges. First, the solar array is significantly larger than the base of the rover; as the array unfurls, the center of gravity eventually sits far above the rover's base. This means that if the array is not perpendicular (relative to the plum vector on the Moon), it may buckle, and worst of all, the entire system is susceptible to tipping over. The operational window for the deployment of the array is thus very precise and any amount of leaning must be carefully tracked and adjusted for (or the deployment process must be cancelled and tried again). A requirement for the unfurling to even begin is the settling of the wheels of the main rover body into the lunar regolith. The VSAT platform will do this by wiggling its wheels (small motions in opposition to each other and opposite to that of the opposing wheel) to dig into the lunar regolith and make the platform more stable for the solar array deployment.


Deployed

10. FAULT SCENARIOS

There are two main types of faults and several ways that they may arise. The most catastrophic is, of course, tipping. But a more likely fault is buckling. Both relate to the location of the center of mass of the solar array. In order to be rolled up when stowed, the booms on either side are made up of short, hollow segments with a center elastic running the entire length under constant tension. As these booms, and the solar array, are unrolled, these segments snap into place, fully mated. During stowing operations, the opposite effect occurs, where a torque is applied at the joint between two segments of the next section that is being rolled up at the bottom of the stowing array in order to unsnap it. Given the amount of tension of the center elastic and the dimensions of each segment, the amount of this torque needed can be easily calculated. This same value is what would cause the array to buckle, most likely at the joint closest to the base, and collapse to the lunar surface, if the array were too far from perfectly vertical. Even more extreme out-of-vertical angles could cause the whole VSAT rover to tip over. A static value has been calculated to prevent these occurrences, and the recommended tolerance is to stay with 3 degrees of vertical. Most of the weight (because that is where the solar cells actually are) is in the upper half of the 60' mast. Combined with the fact that the boom segments act like a spring, it causes the mast to sag further in the direction out of vertical. There is also a potential for an inverted pendulum effect. This could be quite complex because the mast is less stiff in the direction normal to the surface of the mast and more stiff in the direction along the surface of the mast, from one beam toward the other. The inverted pendulum can therefore have an elliptical motion. The inverted pendulum effect, if it occurs, decreases the effective safety tolerance to less than the static value of 3 degrees.

These failure scenarios can be a result of several causes. The sensors used to determine "verticalness" may have too much error. Energy could be added to the system, causing the inverted pendulum, possibly from rough gears or motors or other damage due to launch vibrations. If the soil gives way during deployment, it would introduce both an off-angle *and* energy for the inverted pendulum. The soil could give way as a result of vibrations during mast deployment, the very slow Sun-tracking process, or during unwinding every 28 days (because the solar array cannot continue tracking in one direction more than one full revolution).

11. VSAT MODEL

Several kinds of sensors are onboard the VSAT to help detect anomalies before they become catastrophic. Unfortunately given the round-trip communications delays, it would not always be possible for humans on the ground to take effective action in time, such that the VSAT needs an onboard, autonomous, sense and react capability. The types of sensors include load sensors at each of the wheels, inclinometers, gimbal joint angle sensors, mast deployment motor sensors, and up-facing camera. The data these sensors collect are processed with both the MBR system and TRIAD, which also includes the ML techniques discussed earlier. A high-level architecture of MAIFLOWER applied to VSAT is shown to the right, followed by the model of the mechanical system components.

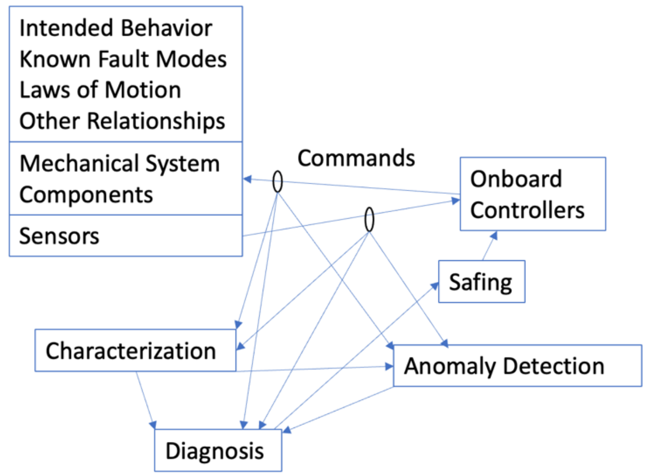


Figure 2: High-Level Architecture of MAIFLOWER Applied to VSAT

Relationships are used to both sanity-check sensor values (i.e., identify faulty sensors) and to diagnose causes of anomalies. An example of a relationship is that using the angle of the body from vertical, the angle of the platform and solar array base, and the state of the solar array (how far deployed, deployment motion, inverted pendulum motion, etc.), the wheel loads can be calculated and cross-checked against the wheel load sensors. Another example is how the vertical motion of the IMU at the top of the mast can be checked against both the up-facing camera and motor sensors which can be used to also calculate the degree of deployment. Cross-vertical motions of the top IMU (i.e., inverted pendulum motions) can be cross- checked against the up-facing cameras. By detecting and diagnosing problems early, before they become so severe that disaster strikes, MAIFLOWER can greatly decrease mission risk. The primary safing mechanism is to stop the current operation and retract the mast (to improve its center of gravity). For example, if during deployment, inverted pendulum motion is detected, stopping all motion would tend to decrease the problem. If there is a danger of tipping, retracting the mast will improve the problem. If the soil starts collapsing or sliding during deployment, solar tracking, or unwinding, halting the motion will stop the vibration.

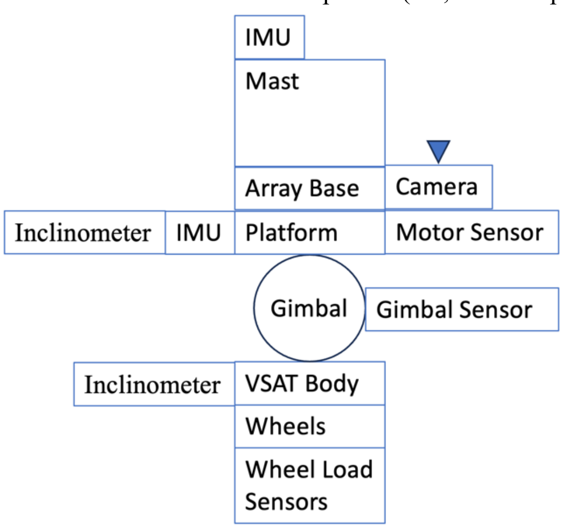
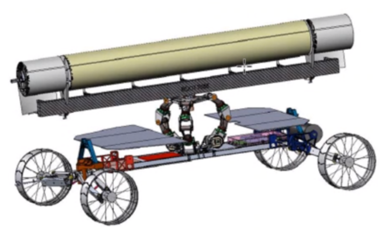


Figure 3: Model of Mechanical System



Driving Configuration

12. CURRENT/FUTURE WORK

Currently we are prototyping MAIFLOWER for VSAT and applying its MBR, TRIAD, and ML techniques to simulated data to detect and diagnose anomalies. A ground unit of VSAT will be tested in simulated lunar regolith in a large vacuum chamber this coming summer and we will then be applying MAIFLOWER to the actual data from the actual vacuum chamber tests with the real ground article. At that time, we will also begin investigating autonomous safing actions.

We hope to further test and mature the system for deployment, with VSAT, to the Moon's surface. Additionally, there is a similar mission planned with another very tall, deployable solar array that will be deployed directly from a lander. Many of the components and possible problems are identical or nearly so to the ones described here for VSAT, so there is potential for us to be involved in that mission as well.

13. CONCLUSIONS

Although, as described earlier, these technologies have been applied to a wide array of spacecraft subsystems, this is the first involving a primarily mechanical system. So far, the process has been similar to our other previous work. This combination of technologies and our current effort present the following benefits:

- A generalized and modular fault management architecture that can be quickly spun up for any number of different subsystems.
- Autonomous, high-speed anomaly detection along with "root cause" analysis by correlating time-series data across subsystems, thereby capturing cascading impacts of single faults on a spacecraft as a whole.
- Introduction of transformer-based anomaly detection to fault detection in the space domain.
- Astrobotic's VSAT, which provides a real platform, first in vacuum chamber ground testing, then on the lunar surface, to prove MAIFLOWER's feasibility for adaptation for other spacecraft.

14. REFERENCES

[1] Gao, H. et al. (2022). Tsmae: A novel anomaly detection approach for Internet of Things time series data using memory-augmented autoencoder. *IEEE Transactions on Network Science and Engineering*.
[2] Geneva, N., & Zabaras, N. (2022). Transformers for modeling physical systems. *Neural Networks*, *146*, 272-289.
[3] Li, Z. et al. (2018). Data-driven bearing fault identification using improved hidden Markov model and self-organizing map. *Computers & Industrial Engineering*, *116*(2018), 37-46.
[4] Ramachandran, S., Rosengarten, M., & Belardi, C. (2020). Semi-supervised machine learning for spacecraft anomaly detection & diagnosis. In *2020 IEEE Aerospace Conference*. IEEE.
[5] Ren, H. et al. (2019). Time-series anomaly detection service at Microsoft. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*.
[6] Smith, D., Frank, J., & Cushing, W. (2008). *The ANML language*. ICAPS-08 Poster Session.
[7] Stottler, R., Ramachandran, S., Healy, C., Singhal, A., & Finnigan, E. (2022). On-board, autonomous, hybrid spacecraft subsystem fault and anomaly detection, diagnosis, root cause determination, and recovery. In *Proceedings of Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS), 2022*.
[8] Stottler, R. (2020). *On-board, autonomous, hybrid spacecraft subsystem fault and anomaly detection, diagnosis, and recovery*. Space Domain Awareness Data Science Working Group, November 2020.
[9] Stottler, R., Ramachandran, S., Belardi, C., & Mandayam, R. (2020). On-board, autonomous, hybrid spacecraft subsystem fault and anomaly detection, diagnosis, and recovery. In *Proceedings of Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS), 2020*.
[10] Xu, J. et al. (2021). *Anomaly transformer: Time series anomaly detection with association discrepancy.* arXiv preprint arXiv:2110.02642.