Planetary Drilling Autonomy

Sarah Boelter[1]

***Abstract*— In planetary environments, robotic agents must reason through faults before they escalate to mission-critical failures. No broadly applicable solution exists to give a specialized agent like The Regolith and Ice Drill for Exploring New Terrain (TRIDENT) awareness for when a situation may escalate to a drilling fault. We propose building on our previous work with online time-series subspace analysis methods and percussive beat frequency detection techniques to define and solve a Markov Decision Process to autonomously avoid drilling faults.**

I. INTRODUCTION

This work focuses on a robotic drill, known as "The Regolith and Ice Drill for Exploring New Terrain" (TRIDENT) [1]. It is a 1-meter rotary percussive drill manufactured by Honeybee Robotics. We collected 30 instances of fault data with the TRIDENT drill with NASA Ames and the Goddard Institute Field Team (GIFT) in the Bishop Tuff in Fall 2023 [2] and at a Haughton Crater Analog site [3] in 2024. TRIDENT gives indicators, visually or otherwise, of a failure well before it happens, but the necessary components to intelligently monitor actions do not exist yet [4][5][6]. We propose building on previous work with online time-series subspace analysis methods and percussive beat frequency detection to develop a Markov Decision Process to autonomously predict and avoid drilling faults. This work is conceptual, and focuses on analysis and developing theoretical solutions for implementation during summer lab testing.

II. PREVIOUS WORK

The Atacama Rover Astrobiology Drilling Studies (ARADS) project [1] aimed to understand the subsurface down to 1 meter in planetary environments using a rover with an attached TRIDENT drill. This system was tested in Chile in 2019. While The ARADS project prototyped task automation, the drill still requires constant oversight. Glass et al. [1] detail how it is impossible to know without prior geological surveying what exists below the surface, especially in planetary environments.

Previous work predicting situational risk factors for critical errors for TRIDENT used change point analysis techniques [8]. Subspace-based methods are best at detecting instantaneous changes while not requiring extensive data. Since the decomposition acts as a filter for random signal contributions, it mainly extracts characteristic shapes and ignores noise for the unique properties of our application to drilling. Using the Enhanced Singular Spectrum Transformation (ESST) [9], instead of computing the decomposition and comparing the extracted characteristics separately, ESST combines both trajectory matrices in one matrix and computes the score directly after decomposition, using more characteristics than the SST [10]. We conducted testing, and saw change score severity can help determine the drill is entering a faulty state.

Techniques from the field of structural health management (SHM) have been used on previous models of extraterrestrial drills [11]. Previous work excelled at detecting oncoming faults, but diagnostic rates were too slow to catch those that onset in under 20 seconds, and were not designed for percussing drills. Our current work builds on previous SHM research by providing an unsupervised ensemble-learning method to identify percussive beats in drill vibration, showing the frequencies present during percussive strikes can be used to perform high-frequency health diagnostics. These frequencies can be coupled with techniques like ESST to help inform machine learning models of oncoming drill faults.

Time-series examples from two different boreholes are in Figure 1. The first shows normal drilling operation and the second illustrates a fault. Faults tend to raise variables like Torque and Weight On Bit (WOB) while Rate of Penetration (ROP) drops to zero. The data is non-static over time, and faults can span as little as a few seconds to several minutes.

III. PROBLEM STATEMENT

We built off previous work of Cayeux et al. [12] to optimize the time needed to drill to the total desired depth. We can decompose this problem into two parts: the time to execute the series of actions, and the time to mitigate any drilling faults if they occur, which can be expressed as:

$$\Delta t_{TD}(A) = \sum_{a \in A} t_a(a) + \sum_{f \in F} P(f) t_m(f) \tag{1}$$

where Δt_{TD} is the estimated time to reach the total depth (TD), A is a set of actions a, t_a is a function that estimates the duration of an action, f is a possible drilling fault from the set of possible drilling faults F, $P(f)$ is the occurrence probability of a drilling fault, and t_m is the estimated time to mitigate a drilling fault. We want to minimize the number of actions to reach the total depth. The objective function is:

$$\underset{A \in \mathcal{A}}{\operatorname{argmin}} \ \Delta t_{TD}(A) \tag{2}$$

where $\mathcal{A}$ is a set of possible actions to reach the TD.

A. *Defining the MDP*

We can represent our MDP for TRIDENT as a five-tuple $\langle S, A, T, R, \gamma \rangle$ where S is a state as defined by equation 3, A is a set of all possible actions in S, which, for example, is

[1]Sarah Boelter is with the Department of Computer Science and Engineering at the University of Minnesota in Minneapolis, MN. boelt072@umn.edu

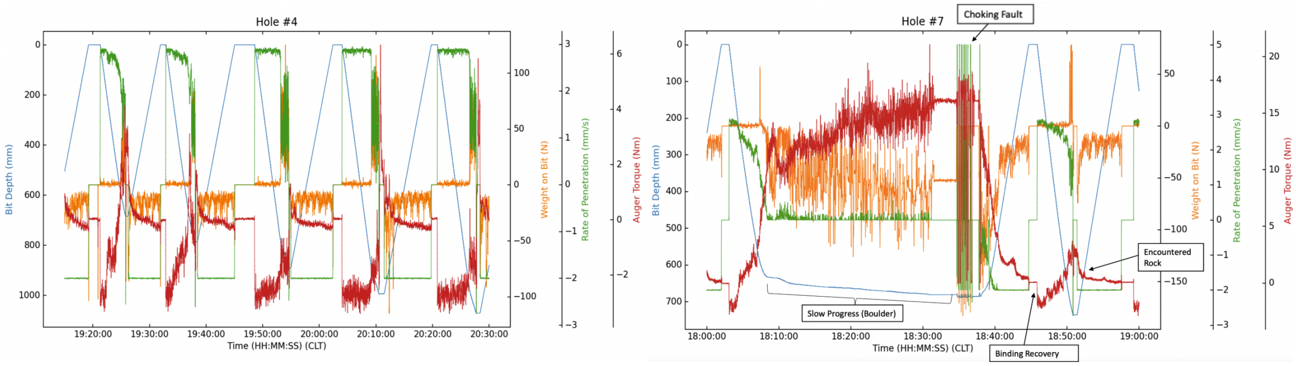

Fig. 1. Left image shows normal drilling operations and right image shows choking fault in Devon Island, Canada [7]

defined as $A = \{continue, retract\}$, where *continue* equates with continuing to drill, and *retract* with retracting before drilling specified depth. T is the transition model, which represents for each state s and action a the probability of reaching another state s'. We write T as $T(s'|s,a)$. R is defined as a reward function $R(s,a)$, where R represents the reward for taking action a in state s. The discount factor γ is used to favor a more immediate reward in the next state versus a more distant reward several states in the future.

We make similar state considerations as Cayeux et al. [12], but must consider fault detection techniques available to us for this hardware. We can define our state as a combination:

$$s = (C_T, C_W, R, P, B_{skip}, D) \tag{3}$$

C_T and C_W are the normalized torque and WOB change scores computed by ESST, R is ROP, P is a binary variable indicating percussion, B_{skip} states if skipped percussive beat was detected, D is depth in meters. We can monitor C_T and C_W, which spike to indicate faults. R can monitor stalled velocity. P indicates hard material composition with active percussion. If percussion is active, we begin monitoring B_{skip}. B_{skip} represents training from beat detection methods. These calculate a novelty curve corresponding to an increase in spectral energy of the vibration signal. Then, an algorithm is used to pick local minima in the novelty curve corresponding to beats. After applying the Fourier transform to beats, features like frequency peaks and overall energy are used to determine if faulty or nominal operation. Test data shows that non-faulting beats typically have a frequency peak around 6 kHz, and faulting beats tend to have higher frequency peaks and more overall energy.

For transition functions, we must understand how variables change under different drilling depths. Since we have extensive non-faulty time-series fieldwork data for analog drilling conditions, we can accumulate statistics for a Gaussian distribution and compute the mean and standard deviation of those data points to come up with the necessary transition functions for incremental depth ranges.

B. Solving the MDP

To solve this MDP, we will use a reinforcement learning framework. Our goal is to find an optimal policy

$$\pi^* = \arg\min_{\pi} \mathbb{E}\left[\Delta t_{TD}(A) \mid \pi\right] \tag{4}$$

By minimizing Δt_{TD}, we minimize the total actions and possible faults. We parameterize a stochastic policy $\pi(a \mid s; \theta)$ using a neural network outputting the probability of selecting each action given the current state s. The agent interacts with its environment where at each timestep it selects an action $a \sim \pi(a \mid s; \theta)$, receives a reward $R(s,a)$, which combines a term for depth progress with a time and fault-penalty term and is defined as

$$R(s,a) = w_1 \, \Delta D - w_2 \left[t_a(a) + \sum_{f \in F} P(f \mid s,a) \, t_m(f) \right] \tag{5}$$

where ΔD is the incremental depth of the drilled, and w_1 and w_2 are the weights for the reward. In space drilling, where faults are extremely costly, w_2 should be set relatively high to ensure that any action that increases fault risk is heavily penalized and the time penalty incurred from a potential drilling event is t_m. We aim to maximize the total drilled depth by accumulating incremental depths ΔD, while minimizing the total time spent, represented by action durations $t_a(a)$. Additionally, $P(f \mid s,a)$ is the probability of encountering a fault f, e.g., choking, binding, over-torque, etc. We will build on Cayeux et al. and use the Gauss-Seidel Value Iteration method [12][13], to converge to the optimal value function:

$$V(s) \leftarrow \max_a \left[R(s,a) + \gamma \sum_{s'} P(s' \mid s,a) V(s') \right], \tag{6}$$

where $V(s)$ is the value function for state s, $R(s,a)$ is the immediate reward received by taking action a in state s, $\gamma \in [0,1]$ is the discount factor for future rewards, $P(s' \mid s,a)$ is the probability of transitioning to state s' given the current state s and action a. We only have to keep one copy of state values in memory instead of two because the values are updated in place. When variables within the current state veer into uncertainty, we use the optimal policy π^* to determine the best course of actions for the current state.

IV. Conclusions and Future Work

Future work will include the implementation and testing of our MDP on the TRIDENT drill during in-lab testing in April and May of 2025. After tuning our solution, we will proceed with testing our lab-perfected methods during analog field testing in Alaska in September 2025.

REFERENCES

[1] B. Glass, D. Bergman, V. Parro, L. Kobayashi, C. Stoker, R. Quinn, A. Davila, P. Willis, W. Brinckerhoff, K. Warren-Rhodes, M. Wilhelm, L. Caceres, J. DiRuggiero, K. Zacny, M. Moreno-Paz, A. Dave, S. Seitz, A. Grubisic, M. Castillo, R. Bonaccorsi, and the ARADS Team, "The Atacama Rover Astrobiology Drilling Studies (ARADS) Project," *Astrobiology*, vol. 23, pp. 1245–1258, Dec. 2023.

[2] B. Glass, C. Stoker, H. Battah, S. Boelter, C.Fortuin, I. King, T. Stevenson, and T. Stucky, "TRIDENT drill validation at Mars Analog Field Sites," in *55th Lunar and Planetary Science Conference (LPSC)*, 2024.

[3] Osinski *et al.*, "Geological overview and cratering model for the Haughton impact structure, Devon Island, Canadian High Arctic," *Meteoritics & Planetary Science*, vol. 40, no. 12, pp. 1759–1776, 2005. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1945-5100.2005.tb00145.x

[4] N. Muscettola, P. P. Nayak, B. Pell, and B. C. Williams, "Remote agent: to boldly go where no AI system has gone before," *Artificial Intelligence*, vol. 103, no. 1–2, pp. 5–47, Aug. 1998. [Online]. Available: https://doi.org/10.1016/S0004-3702(98)00068-X

[5] G. Paulsen, K. Zacny, P. Chu, E. Mumm, K. Davis, S. Frader-Thompson, K. Petrich, D. Glaser, P. Bartlett, H. Cannon, and B. Glass, "Robotic drill systems for planetary exploration," in *Space 2006*.

[6] J. Wardt, E. Cayeux, R. Mihai, J. Macpherson, P. Annaiyappa, and D. Pirovolou, "Taxonomy describing levels of autonomous drilling systems: Incorporating complexity, uncertainty, sparse data, with human interaction," in *IADC/SPE International Drilling Conference and Exhibition*, 2024.

[7] B. Glass, C. Fortuin, H. Battah, and I. King, "TRIDENT Drill Validation Testing in Haughton Crater, Devon Island, Canada," ASCE Earth and Space 2024, 2024.

[8] S. Boelter, E. Temesgen, B. Glass, and M. Gini, "Understanding drill data for autonomous application," in *International Workshop on Autonomous Agents and Multi-Agent Systems for Space Applications at AAMAS*, 2024.

[9] L. Weber, ""changepoynt"," 2024. [Online]. Available: https://github.com/Lucew/changepoynt

[10] T. Idé and K. Inoue, "Knowledge discovery from heterogeneous dynamic systems using change-point correlations," in *Proceedings of SIAM International Conference on Data Mining*, 2005, pp. 571–575.

[11] S. Statham, "Autonomous structural health monitoring technique for interplanetary drilling applications using laser doppler velocimeters," Ph.D. dissertation, Georgia Institute of Technology, 2011.

[12] E. Cayeux, B. Daireaux, A. Ambrus, R. Mihai, and L. Carlsen, "Autonomous decision-making while drilling," *Energies*, vol. 14, p. 969, 2021.

[13] M. J. Kochenderfer, C. Amato, G. Chowdhary, J. P. How, H. J. D. Reynolds, J. R. Thornton, P. A. Torres-Carrasquillo, N. K. Üre, and J. Vian, *Decision Making Under Uncertainty: Theory and Application*, 1st ed. The MIT Press, 2015.