Doppler-Based Geolocation with a Single LEO Satellite: A Machine Learning Perspective

Octave Dupuis
Eutelsat
Issy-les-Moulineaux, France
odupuis@eutelsat.com

Florian Collard
Eutelsat
Issy-les-Moulineaux, France
fcollard@eutelsat.com

Geoffroy Brichler
Eutelsat
Issy-les-Moulineaux, France
gbrichler@eutelsat.com

Hugo Garny
Eutelsat
Issy-les-Moulineaux, France
hgarny@eutelsat.com

Soufiane Zekri
Commandement de l'Espace
Toulouse, France
soufiane.zekri@intradef.gouv.fr

Léo Dubois-Dunilac
École de l'Air et de l'Espace
Salon-de-Provence, France
leo.dubois-dunilac@ecole-air.fr

***Abstract*—This paper introduces and presents results of a geolocation algorithm (at the level of a LEO satellite) of a transmitter (on Earth) using Doppler estimation. Although the method is well known, the main feature of this algorithm lies in its reliance on machine learning (ML) techniques. A large set of theoretical simulations demonstrate sub-kilometer accuracies even under limiting conditions: a single satellite in sight, a short signal acquisition time, and fluctuating on-board electronics, both in receiver and emitter side. The results highlight correlations between some of the system parameters (satellite pass geometry, frequency acquisition, signal quality) and geolocation accuracy. The algorithm has been validated on real data using a CubeSat, demonstrating performances in operational conditions very close to the theoretical ones.**

***Index Terms*—Geolocation, Machine Learning (ML), Low Earth Orbit (LEO), Doppler Effect**

I. Introduction

In the past decade, the space environment of the Low Earth Orbit (LEO) has become increasingly dense, a consequence of the massive deployment of satellites, especially constellations such as Starlink, Guowang or OneWeb. This densification of satellites led to the emergence of new telecommunications services, as well as an increasing number of signals transmitted from Earth to Space. In this context, the use of inherent characteristics of LEO satellites to locate a signal source from Earth is a common area of interest, through two major stakes:

- Locate a known emitter that does not have a GNSS receiver (for low-cost and low-consumption applications) or whose GNSS receiver is jammed.
- Locate an unknown emitter that interferes with a signal of interest.

Throughout this paper, we focus on a geolocation solution to cope with the two major stakes above under very limiting conditions: a single satellite in sight, a single pass over the emitter, a short signal acquisition time, and volatile signal processing parameters linked to on-board electronics.

There are two different ways of achieving geolocation in this context: by exploiting the Angle of Arrival (AoA) of the signal during the passage of the satellite, or its Frequency of Arrival (FoA), which is subject to a Doppler shift. Our geolocation technique is based on FoA, which is technically better suited to our case.

As the satellite moves relative to the transmitter, the signal undergoes a frequency drift: this is the Doppler effect that allows a position to be estimated. This process is well known and has been the subject of numerous articles. However, although the theory of Doppler geolocation is well documented, its operational implementation adds numerous constraints, particularly in terms of data accuracy (temporal and frequency resolution, frequency tracking method, RF impairments, etc.).

This study builds on a diverse body of research on Doppler-based geolocation and single-satellite positioning methods. In particular, Nguyen et al. introduced algebraic approaches for initializing Doppler-based geolocation algorithms, reducing computational complexity while achieving reliable convergence [1]. Ellis et al. explored the use of constrained Unscented Kalman Filters (cUKF), incorporating sigma point projection techniques to improve geolocation accuracy on Earth's surface under real-time constraints [2,3,4]. In addition, Lopez et al. [5] enhanced Argos Doppler positioning with multiple-model Kalman filtering, significantly improving accuracy in challenging scenarios with sparse data by leveraging state-space models for trajectory refinement.

In this article, we will briefly introduce our machine learning based geolocation algorithm, referenced EOTT (Eye Of The Tiger), before delving deeper into the test campaigns designed to evaluate its performance.

II. Algorithm description

A. Generic scheme

EOTT is powered by Support Vector Regressions (SVR) with radial basis function kernels to predict a position based on three input parameters:

- A few frequency estimation measurements of the signal received by the satellite (referred simply as doppler measurements in the following).
- The satellite ephemeris.
- The signal transmission frequency.

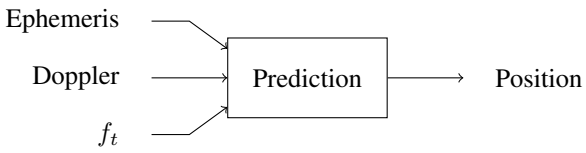

Fig. 1. EOTT description

Being based on supervised ML techniques, EOTT requires a dataset on which to train i.e. a set of ephemerides, dopplers at a certain frequency, and associated positions. Our approach was to create a synthetic one, in line with our use case. We generate simulated ephemeris corresponding to a certain LEO orbit (depending of our usecase). We then chose sets of coordinates from which to calculate the associated Doppler curves, using the following theoretical formula:

$$f_d = -\frac{f_c}{c}.(v_s - v_t).\frac{p_s - p_t}{\|p_s - p_t\|} + \epsilon$$

In this equation, we have : f_d the Doppler frequency shift between satellite and receiver, c the speed of light, f_c the carrier frequency, v_s, p_s and v_t, p_t are the velocity and position vectors for the satellite and the transmitter respectively, and the global error ϵ.

To generate realistic doppler curves, we statistically model this ϵ term of frequency error. This approach enabled us to model frequency error considering parameters that can be difficult to quantify theoretically - such as biases generated by on-board electronics during signal acquisition or by imperfect modeling of signal propagation.

The two following figures illustrate a sample of our dataset, displaying an orbit and a chosen position from which we calculate a Doppler curve.

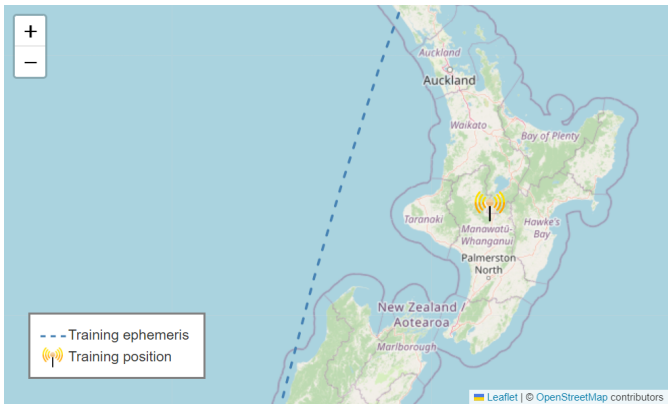

Fig. 2. Example of training ephemeris and position

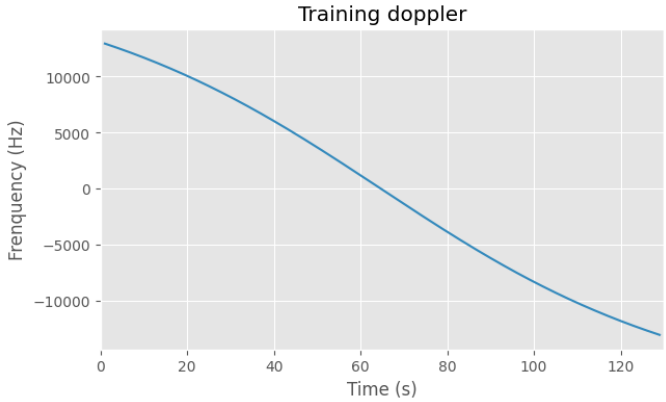

Fig. 3. Example of doppler curve (associated to fig. 2)

Once our dataset created, all that remains is to train and optimize our model. We therefore exploit machine learning to ensure that, in the absence of a rigorously perfect model, we can train ourselves to correct patterns identified on real recordings, without having to know their exact nature.

In the case of geolocation of unknown signals (with the hypothesis that we can obtain frequency estimations), the absence of information on the transmission frequency is overcome by estimating it, which enables us to maintain the same generic algorithm scheme. To do that, we find the minimum of the doppler rate (i.e. derivative).

Let us note that our inputs have inherent parameters: the number of acquisition points in frequencies, and their temporal distribution (one point every 5, 10, 30, etc. seconds).

B. Frequency error modeling

One of the key points in this method is the definition of ϵ. In a first instance, we naturally modeled it as a Gaussian noise, with a defined standard deviation (std). However, after statistics analysis performed on real transmissions, we figured that it was not exactly Gaussian. We therefore carried out measurements of IoT device transmissions using a CubeSat payload laboratory replica and defined a "custom noise" with the same distribution as the observed errors (still defined by a std). This modeling enables us to better analyze the biases and sources of error generated by on-board electronics. We note that one major difference with the Gaussian noise is that the custom noise does not have a zero mean.

C. Computational cost

The computational cost of EOTT is primarily driven by the use of Support Vector Regressions (SVRs) with Gaussian kernels. As will be developed later on, we worked with two types of data acquisition: 5 or 121 frequency measurements. The total computational requirement can be summarized as follows (theoretically):

- Preset A (5 measurements) $\sim 2.10^5$ FLOPs.
- Preset B (121 measurements) $\sim 1.10^6$ FLOPs.

These values ensure that the algorithm is computationally efficient and suitable for deployment on resource-constrained platforms such as CubeSats. Further optimizations could reduce this cost even further while maintaining accuracy.

This algorithm, developed and validated with a CubeSat using ISM band frequencies, is now being adapted to other types of LEO satellites, and at any frequency.

III. Theoretical Simulations

A. Methodology

In order to do a proper statistical benchmark on our algorithm, covering many system configurations, we decided to emulate two satellite pass scenarios (which partially correspond to our two over-the-air scenarios, described later on) corresponding to our use case, which means a SSO orbit at 500 km of altitude. For each of these orbits, we chose 900 positions in a square of 1000 km^2, delimiting the sight of a satellite. We computed their doppler at a defined frequency (868.7 MHz), using the two frequency error models presented above, with different stds. A similar process to the creation of the training dataset was used to generate the benchmark dataset. To maintain unbiased results, these two datasets are disjoint. Then, we fed these to our algorithm and calculated the accuracy of the prediction.

As mentioned above, the number of acquisition points in the doppler curve and its temporal distribution are parameters in the algorithm. To match with real operational requirements, we decided to confront two presets :

- Preset A : 5 measurements, one each 30 seconds (corresponding to low powered typical IoT devices).
- Preset B : 121 measurements, one each second (corresponding to classical transmissions).

Moreover, it is set so that the middle measurement finds itself at the moment when the elevation between the transmitter and satellite is at the highest (we determine it by finding the doppler rate's peak).

B. Results

With these theoretical tests, we obtained sub-kilometers accuracies and demonstrated correlations between performance and parameters such as the quality of the transmitted signal or the geometry of the satellite path. Fig. 4 shows a strong correlation (that was expected) between the std of the used noise and the accuracy.

Another correlation observed lays between the accuracy and the maximum of elevation between the transmitter and the satellite (we remind that maximum elevation of a transmitter relative to a satellite is the highest angle, measured from the transmitter's local horizon, at which the satellite appears during its pass). Fig. 5 represents the accuracy as a function of the maximum of elevation for the preset A, with the custom noise at 5 Hz of std.

For more clarity, we represent the results as boxplots, one boxplot gathering the accuracy for positions with a maximum

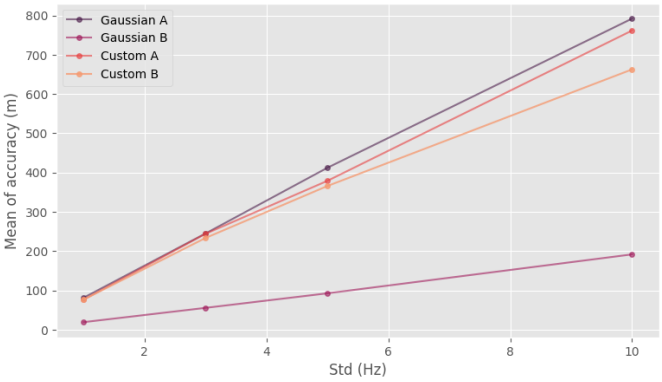

Fig. 4. Mean accuracy as a function of noise standard deviation

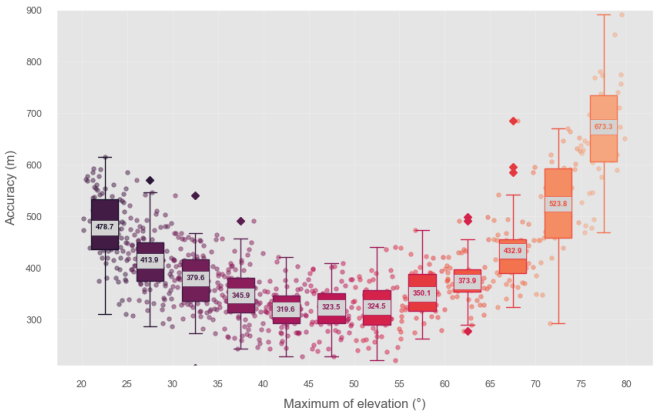

Fig. 5. Accuracy in function of maximum of elevation (preset A, 5 Hz)

of elevation range of 5 degrees. Instead of representing the median, we choose to show the numerical value of the mean for each boxplot. An interesting aspect of this correlation is that with a prediction, we can easily estimate the corresponding maximum of elevation. Thus, with our prediction, we can estimate the accuracy of the prediction.

According to our estimations, the standard deviation of our real data should be 5 Hz. So, here are the numerical values of our theoretical benchmark with both noises at 5 Hz of std :

TABLE I
Theoretical results at 5 Hz of std (accuracies)

Preset	Noise	Mean	Median	Std
A	Gaussian	412.04 m	376.70 m	124.54 m
	Custom	378.77 m	355.70 m	112.33 m
B	Gaussian	92.38 m	85.63 m	29.30 m
	Custom	365.63 m	354.50 m	99.40 m

From these results, it can be observed a relatively significant difference between the two noises used for preset B. This is be because Gaussian noise, with its zero mean, tends to cancel out

due to statistical averaging, thereby reducing its overall impact. In contrast, custom noise with a non-zero mean introduces a consistent bias that degrades performance.

IV. Over the air results

A. Methodology

Finally, we validated our algorithm on real data, using our CubeSat ELO-2 (Eutelsat LEO for Objects) and two antennas located in France and in Brazil, emulating typical low-power signal from an IoT device at frequencies of 868.7 MHz and 902.3 MHz during multiple satellite passes (36 in France, 26 in Brazil).

With reference values, we can look at real data, and quantify the performance of our algorithm.

B. Results

The results obtained on the field are fairly close to those obtained from theoretical simulations with sub-kilometers accuracies, demonstrating the robustness of the algorithm under real-life conditions. In order to see the impact of the different types of noises, we computed the results for both of them:

TABLE II
Real-life results (accuracies)

Preset	Noise	Mean	Median	Std
A	Gaussian	672.85 m	550.56 m	509.08 m
	Custom	468.08 m	426.76 m	358.58 m
B	Gaussian	650.47 m	535.16 m	469.10 m
	Custom	441.43 m	392.47 m	288.99 m

We observe that the results using Gaussian noise training are less good than those with our custom noise, indicating that this more realistic error modeling improves the performances of our algorithm. The main difference lies in the stds between theoretical and real results with the custom noise. We can assume that this is due to the still imperfect model of noise and errors (clock drift, atmospheric effects, etc.).

The next two figures show the maps with the predictions and the actual positions of the transmitters for preset A, with custom noise set at 5 Hz. The predicted positions, if averaged, tend to converge on the actual positions.

V. Conclusion

This work demonstrates the effectiveness of a Doppler-based geolocation algorithm using machine learning techniques. Validated by simulations and tests on real data with a CubeSat, the algorithm achieves sub-kilometric accuracies despite severe constraints (short acquisition time, fluctuating onboard electronics). The approach, centered on an SVR model and a synthetic dataset, reveals key correlations between orbit geometry, signal quality and accuracy. These promising results pave the way for robust practical applications in complex scenarios and contribute to advancing geolocation capabilities for LEO satellite systems.

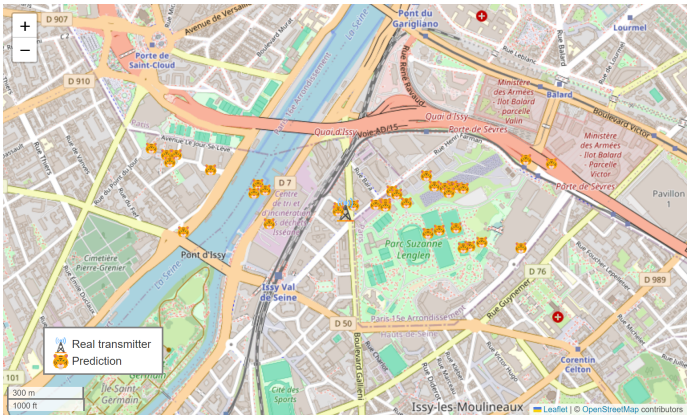

Fig. 6. Predictions in France

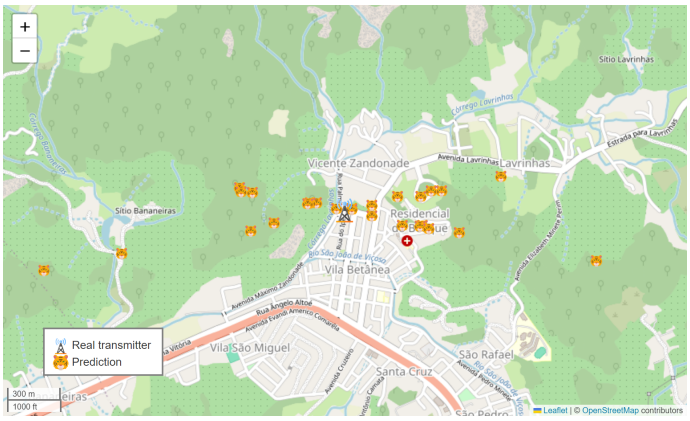

Fig. 7. Predictions in Brazil

References

[1] Nguyen, N. H., & Doğançay, K. (2016, March). Algebraic solution for stationary emitter geolocation by a LEO satellite using Doppler frequency measurements. In 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 3341-3345). IEEE.

[2] Ellis, P. B. (2020). Geolocation of a Radio Frequency Emitter Using a Single Low Earth Satellite. University of California, Santa Cruz.

[3] Ellis, P., & Dowla, F. (2018, June). A single satellite geolocation solution of an RF emitter using a constrained unscented Kalman filter. In 2018 IEEE Statistical Signal Processing Workshop (SSP) (pp. 643-647). IEEE.

[4] Ellis, P., & Dowla, F. (2018, September). Performance bounds of a single LEO satellite providing geolocation of an RF emitter. In 2018 9th Advanced Satellite Multimedia Systems Conference and the 15th Signal Processing for Space Communications Workshop (ASMS/SPSC) (pp. 1-5). IEEE.

[5] R. Lopez, J. -P. Malardé, F. Royer and P. Gaspar, "Improving Argos Doppler Location Using Multiple-Model Kalman Filtering," in IEEE Transactions on Geoscience and Remote Sensing, vol. 52, no. 8, pp. 4744-4755, Aug. 2014, doi: 10.1109/TGRS.2013.2284293.

[6] YU, Kegen. Positioning and Navigation Using Machine Learning Methods. 2024.

[7] Andreotti, R., Andrenacci, M., & Nanna, L. Radio Frequency Signals Detection and Geolocation using single LEO satellite in the VHF band.