Hysteresis in RAIM

Juan Blanch, Christoph Mayer*, Sherman Lo, Todd Walter, Per Enge.
Stanford University
**German Aerospace Center*

ABSTRACT

Absolute RAIM algorithms proposed for vertical guidance are based on a snapshot approach. At a given time, each satellite has a probability of being in a faulted state which does not depend on how long the satellite has been in view. This approach greatly simplifies the proof of safety. Both the threat model and the algorithm are simple enough so that the Probability of Hazardously Misleading Information can be shown to be below the required level for vertical guidance, even in the multiple failure case. However, this approach neglects the observation history of the satellites. A faulty satellite that has been in view for several hours has more chances to be detected and excluded from the position solution than a satellite that has just risen. This distinction is absent in the snapshot approach, therefore ignoring useful information –even though snapshot RAIM will have hysteresis in the exclusion function. In the Absolute RAIM (ARAIM) architecture proposed in [1] it is assumed that a failure could have started at any point in a defined period of time, which is unrelated to the time the satellite has been in view for the user. This period of time is the time it will take the GNSS provider to either flag the satellite or correct it.

The goal of this paper is to define a threat model and a user algorithm that takes into account the satellite observation history so that performance can be increased. The threat model must be such that it includes all reasonably possible failures, without being overly optimistic – this is very important as even small deviations from the nominal behavior can be problematic for vertical guidance. The introduction of the time aspect in the threat model makes it much more complicated than in the snapshot approach. The failure of one satellite needs to be divided into several possibilities depending on when the failure started. Even within a given satellite failure at a given time, the failure is an error profile, not a given scalar (like in snapshot RAIM). All of this is also complicated by the fact that nominal errors are correlated in time. The algorithm must be practical, and most importantly, must allow a clear evaluation of the Probability of Hazardously Misleading Information. We propose a threat model that conservatively describes the possible threats in the time domain and an algorithm that mitigates this threat model. Once the algorithm is defined we will evaluate the gain in performance for worldwide vertical guidance.

INTRODUCTION

With dual frequency (L1-L5), the possibility of using Receiver Autonomous Integrity Monitoring for vertical guidance is very attractive [1]. However, an issue with RAIM for vertical guidance is that the expected L1-L5 constellation might not be sufficient for some applications, in particular for civil aviation. Past RAIM coverage analyses have shown that a 24 satellite constellation has a low coverage of high LPV 200 availability. These analyses have relied on a snapshot approach, that is, the error bound computed at a given time only relies on the range measurements done at that time. This approach has the advantage of simplifying the error bound computation and the error analysis. However, it has the disadvantage of ignoring useful information in the calculation of the error bound. In snapshot RAIM it is assumed that a failure could have started at any point in the previous hours and would have remained undetected until the current time.

In a real setting, a satellite that has caused a test to trip will be excluded temporarily. (This time will be called the hysteresis time in this paper.) It is therefore pessimistic to assume (as snapshot RAIM does) that the a priori probability of failure is the probability that the failure started at any point in the previous hours. The objective of this paper is to evaluate the performance gained by taking into account the history of the measurements, as doing this can only lower the error bounds. The analysis needs to be both conservative and feasible. The outline of the paper is as follows: first we will describe the threat model, then we will describe the algorithm as a simple generalization of a solution separation RAIM, we will present the results of the availability simulations, then we will suggest a methodology to account for the correlation of the nominal pseudorange range errors.

THREAT MODEL

In snapshot RAIM, a failure is a bias in a given satellite (or set of satellites if one considers multiple failures) at the time of interest t_0. When the history of the fault is considered, one needs to consider the whole error profile. The history of the fault until t_0 is relevant, as it affects the probability that the failure might have been detected before t_0. For the purpose of this discussion we will label $b(t)$ the error profile and b_0 – the bias that affects the user position – is $b(t_0)$.

The most conservative threat model that can be considered is one where all threats remain well below the noise level for all times before t_0 and then rise to an arbitrary $b(t_0)$. This threat model is the most comprehensive, however it ignores the fact that faults develop over time. This is in fact the threat model that is assumed in the snapshot RAIM approach. Since here we are trying to exploit the temporal characteristics of the threats we need a threat model that keeps a more realistic relationship between b_0, the value at t_0 and $b(t)$. The threat model that is proposed here is based on the following observation. For any error profile $b(t)$, there exists a ramp error $r(t)$ such that:

$$
\begin{aligned}
r(t) &\leq |b(t)| \\
r(t_0) &= |b_0|
\end{aligned}
\tag{1}
$$

and that overbounds all other ramps. The exact expression for such a ramp is:

$$
\begin{aligned}
r(t) &= \max\left(\alpha\left(t-t_0\right)+b_0,0\right) \\
\alpha &= \max\left(\frac{b(t)-b_0}{t-t_0}\right)
\end{aligned}
\tag{2}
$$

The important point is that this ramp has a larger impact on integrity than $b(t)$, because it is less detectable in the past, but has the same magnitude at t_0. This means that any error profile can be conservatively represented by a ramp error. Let t_{start} be the starting point of the ramp.

The assumption made in this paper is that the ramps obtained from the actual error profiles have a starting point t_{start} that is equiprobable in the interval [t_0-TTA t_0], where TTA is the time to alert from the ground, which is the maximum time a faulted state can persist on a satellite without warning to the user from the ground. It is reminded here that the Absolute RAIM architectures contemplated in [1] all assume that satellite failures cannot persist indefinitely in a satellite.

At this stage, the threat model is a collection of ramps starting between t_0-TTA and t_0. To simplify it further, we divide this interval into equal subintervals ΔT and model all ramps starting between t- ΔT and t by a ramp starting at t. This is a conservative step because we are assuming that the ramps start later than in the continuous model.

Nominal Pseudorange Error

The behavior of the nominal error also needs to be described in the threat model. For the purposes of this paper we will assume that the errors are zero mean Gaussian using the model described in [1]. The URA value assumed will be 1 m. Because the history of the measurements is going to be used, it is necessary to describe the temporal correlation of the errors. In the algorithm described in the next paragraph, it will be assumed that the error is decorrelated after two minutes, which corresponds approximately to the carrier smoothing time constant of 100 s. This is a simplification, because the clock and ephemeris errors, which account for half the error budget, have a correlation time that exceeds two hours. Although this issue will not be solved in this work, we will suggest ways to address it in a later section.

Numerical values for the threat model

A Time to Alert of one hour will be assumed. This means that a failure cannot have started more than one hour before the time t_0. The interval ΔT can be chosen arbitrarily. We will choose here to make it match the two minute decorrelation time.

ALGORITHM

The Absolute RAIM baseline algorithm described in [2] and [1] cannot be easily generalized to the present problem. We therefore describe a variant of a solution separation algorithm. At each step, the solution separation between the all-in-view solution and each subset solution (which should include all possible fault modes) is monitored. If this separation is larger than a given threshold for a given subset, the satellite or set of satellites causing the trip is excluded for the period of time T_{hys}. Snapshot algorithms do not need to specify the time the satellite is excluded (the hysteresis time), because it has no explicit effect on the calculation of the error bound. When the time history of the test statistics is considered we do need to specify the hysteresis time T_{hys}. For this description we will only assume single satellite failures. The method can easily be generalized to multiple failures.

Algorithm without hysteresis

Let L be the error bound for the vertical error. L has to be such that the sum of contributions to the Hazardously Misleading Information (HMI) from each of the failure modes is below the probability of HMI budget (PHMI):

$$\sum_{i=0}^{N_{sat}} P\left(HMI \mid \text{sat. i fails}\right) P_{ap,i} \leq PHMI \quad (3)$$

In this equation N_{sat} is the number of satellites $P_{ap,i}$ is the a priori probability of the failure of satellite i. We have:

$$P\left(HMI \mid \text{sat. i fails}\right) = P\left(\left|\text{solution separation}\right| \leq T,\ \left|\text{position error}\right| > L\right) \quad (4)$$

The expression for the solution separation is given by:

$$\text{solution separation} = s^T \varepsilon + s_i b - s_{subset}^T \varepsilon \quad (5)$$

In this equation s is the set of coefficients that project the measurements onto the position computed using least squares; s_i is the i^{th} component of s. Similarly s_{subset} is the set of coefficients corresponding to the subset where satellite i is excluded. ε is the vector of nominal errors. The position error is given by:

$$\text{position error} = s^T \varepsilon + s_i b \quad (6)$$

For a given bias b we have:

$$\begin{aligned} &P\left(HMI \mid \text{sat. i has bias b}\right) = \\ &P\left(\left|s^T \varepsilon + s_i b - s_{subset}^T \varepsilon\right| \leq T,\ \left|s^T \varepsilon + s_i b\right| > L\right) \end{aligned} \quad (7)$$

The variable b is the magnitude of the fault on satellite i. Because the bias b can be anything, the contribution is obtained by maximizing the above expression. We therefore have:

$$\begin{aligned} &P\left(HMI \mid \text{sat. i fails}\right) = \\ &\max_b \left\{ P\left(\left|s^T \varepsilon + s_i b - s_{subset}^T \varepsilon\right| \leq T,\ \left|s^T \varepsilon + s_i b\right| > L\right) \right\} \end{aligned} \quad (8)$$

It is useful at this point to do the change of variable:

$$u = \frac{s^T \varepsilon - s_{subset}^T \varepsilon}{\sigma_{ss}} \quad (9)$$

$$v = \frac{s^T \varepsilon}{\sigma_v} \quad (10)$$

σ_{ss} is the standard deviation of the solution separation. σ_v is the standard deviation of the nominal error. It is shown in the Appendix that u and v are independent. We therefore have:

$$\begin{aligned} &P\left(HMI \mid \text{sat. i fails}\right) = \\ &\max_b \left\{ P\left(\left|\sigma_{ss} u + s_i b\right| \leq T\right) P\left(\left|\sigma_v v + s_i b\right| > L\right) \right\} \end{aligned} \quad (11)$$

This can be evaluated by finding the bias that maximizes the expression:

$$\begin{aligned} &P\left(\left|\sigma_{ss} u + s_i b\right| \leq T\right) P\left(\left|\sigma_v v + s_i b\right| > L\right) = \\ &\left(Q\left(\frac{T - s_i b}{\sigma_{ss}}\right) - Q\left(\frac{-T - s_i b}{\sigma_{ss}}\right)\right)\left(Q\left(\frac{-L - s_i b}{\sigma_v}\right) + Q\left(\frac{-L + s_i b}{\sigma_v}\right)\right) \end{aligned} \quad (12)$$

In this equation Q is the left sided cdf of a normal distribution. The error bound is found by determining an error bound L such that:

$$\sum_{i=0}^{N_{sat}} P\left(HMI \mid \text{sat. i fails}\right) P_{ap,i} \leq PHMI \quad (13)$$

The threshold T is related to the required probability of false alarm under nominal conditions. The threshold T can be chosen as follows:

$$T = K\sigma_{ss} \quad (14)$$

$$K = Q^{-1}\left(\frac{P_{fa}}{2N_{sat}}\right) \quad (15)$$

Algorithm with hysteresis

The previous algorithm can be generalized to the case with hysteresis simply by accounting for the previous tests:

$$\begin{aligned} &P\left(HMI \mid \text{sat. } i \text{ fails at time } t_{start}\right) = \\ &P\left(\begin{aligned} &\left|\text{solution separation at time } t_0 - T_{hys}\right| \leq T\left(t_0 - T_{hys}\right), \\ &\ldots, \left|\text{solution separation at time } t_0\right| \leq T\left(t_0\right), \\ &\left|\text{position error}\right| > L \end{aligned}\right) \end{aligned} \quad (16)$$

Although a real time receiver would test the solution separation at a higher rate, we only account for the independents tests; these are tests done every ΔT (two minutes). We label t_k the time steps where t_n is the time of the oldest test. We have:

$$t_k = t_0 - k\Delta T \quad (17)$$
$$t_n = t_0 - n\Delta T = t_0 - T_{hys}$$

For a given error profile *b(t)* we have:

$$P\left(HMI \mid \text{sat. i has bias b(t)}\right) = P\begin{pmatrix} \left|s^{(n)T}\varepsilon^{(n)} + s_i^{(n)}b\left(t_n\right) - s_{subset}^{(n)T}\varepsilon^{(n)}\right| \le T_n \\ \vdots \\ \left|s^{(0)T}\varepsilon^{(0)} + s_i^{(0)}b\left(t_n\right) - s_{subset}^{(0)T}\varepsilon^{(0)}\right| \le T_0 \\ \left|s^{(0)T}\varepsilon + s_i^{(0)}b\left(t_0\right)\right| > L \end{pmatrix} \quad (18)$$

The superscript or subscript k indicates the time step at which the variables are considered. Because of the independence of the events, a change of variable gives:

$$P\left(HMI \mid \text{sat. i has bias b(t)}\right) = P\left(\left|{}^{n}\sigma_{ss}u_n + {}^{n}s_i b\left(t_n\right)\right| \le T\right)...P\left(\left|{}^{0}\sigma_{ss}u_0 + {}^{0}s_i b\left(t_0\right)\right| \le T\right) P\left(\left|{}^{0}\sigma_v v + {}^{0}s_i b\left(t_0\right)\right| > L\right) \quad (19)$$

As described in the threat model for each satellite the ramp can start at any of the points t_k. Each of the starting points corresponds to a branch in the threat model. For each starting point we have:

$$b_k\left(t\right) = b\max\left(\frac{t-t_k}{t-t_0}, 0\right) \quad (20)$$

The contribution to the integrity budget is obtained by maximizing the expression (19) over all values of b. Notice that there are $n+1$ contributions per subset. In the single failure case, there are therefore $N_{sat}(n+1)$ branches in the fault tree (to which the fault free case has to be added). To compute the smallest error bound, the following equation in L is solved:

$$\sum_{i=0}^{N_{sat}}\sum_{k=0}^{n} P\left(HMI \mid \text{sat. i fails at } t_k\right)P_{ap,i,k} = PHMI \quad (21)$$

This equation can be solved using interval halving. Notice that for each evaluation and each term of this sum, it is necessary to maximize over b in equation (19). Also, the continuity factor K defined in (15) needs to account for the number of times a satellite could be excluded:

$$K = Q^{-1}\left(\frac{P_{fa}}{2\left(n+1\right)N_{sat}}\right) \quad (22)$$

Increasing the hysteresis time will always decrease the error bound up to TTA, (after that the measurements are no longer relevant to the present time t_0 so the error bound remains the same). On the other hand, increasing hysteresis time can decrease availability, because once a satellite has tripped the test, it remains out of the solution for the period T_{hys}. Also, the computational load increases with the hysteresis time.

PERFORMANCE STUDY

Pseudorange error model

The error model for each error source used here can be found in [1]. It includes the tropospheric error bound, the receiver noise, and the multipath. The User Range Accuracy (URA), which includes the clock and ephemeris errors, was taken to be 1. As opposed to [1] no nominal bias is assumed in this analysis.

Navigation Requirements

The integrity requirement (Probability of Hazardously Misleading Information) is 10^{-7} per approach, and the continuity requirement (Probability that the VPL exceeds the predicted VPL once the approach has started) is 4.10^{-6}. These requirements are very similar to the SBAS requirements. The simulations will evaluate the availability of LPV 200. In these simulations LPV 200 is said to be available if the VPL is below 35 m.

Constellations

The constellations with 24, 27, and 30 satellites are described in [1]. The constellation labeled “current” was obtained by using an almanac from 2008. It is a 24 satellite constellation with several slots containing two satellites. It is supposed that all satellites are dual frequency L1-L5.

Simulation

The Service Volume analysis tool MAAST [3] was used to evaluate the performance of the concept investigated here. The VPL was computed worldwide by simulating a user every 10 degrees in latitude and longitude, for a period of 24 hours every 2 minutes.

Threat Model and Algorithm

The threat model used is the one described above with a ground Time to Alert of one hour. The algorithm used to compute the Vertical Protection Level VPL is the one described in the Algorithm Description section.

Example of effect of hysteresis on VPL

Figure 1 shows the VPL of a user during 12 hours in three cases: the baseline case, 20 minute hysteresis and 1 hour hysteresis. Although the effect appears to be modest for this user, one can see that the only instance of unavailability is removed using hysteresis. It is also interesting to see that a 20 minute hysteresis time provides most of the benefit.

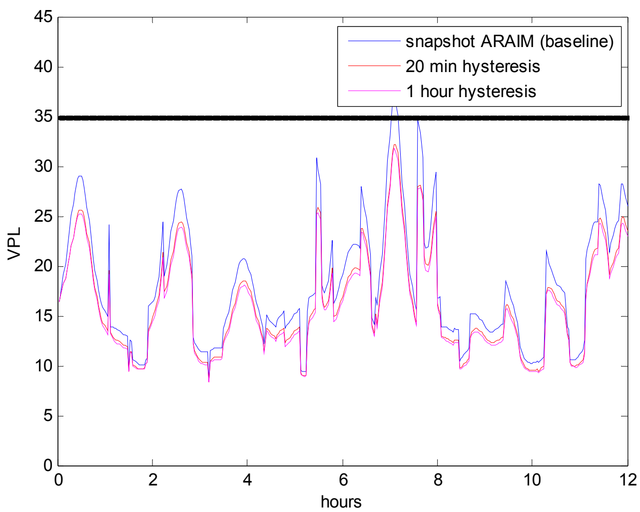

Figure 1. *VPL of a user using different hysteresis times*

Availability results

The availability results are shown in Table 1. The coverage of 99.5% availability of 35 m VPL and 50 m VPL is shown in the first four rows. The fifth and sixth rows show the average 99.5% VPL. The last two rows show the worldwide average availability of 35 m VPL. The use of hysteresis provides a benefit in all the studied constellations. In particular, the coverage of 99.5% availability goes from 80.3% to 91.2% in the "current" constellation, and from 79.9% to 94.3% in the 27 satellite constellation. In the 24 satellite constellation, the average worldwide unavailability is reduced by 50%.

Constellation	Algorithm	24 sat.	Current	27 sat.	30 sat.
Coverage of 99.5% availability of 35 m VPL	Baseline	23.6%	80.3%	79.9%	99.59 %
	1 hour hysteresis	37.2%	91.3%	94.3%	100%
Coverage of 99.5% availability of 50 m VPL	Baseline	63.6%	98.0%	99.5%	99.9%
	1 hour hysteresis	74.6%	98.7%	99.9%	100%
Average 99.5% VPL	Baseline	51.7 m	28.0 m	29.7 m	20.0 m
	1 hour hysteresis	42.6 m	24.7 m	26.3 m	18.0 m
Average Availability of 35 m VPL	Baseline	98.1%	99.9%	99.7%	99.99 %
	1 hour hysteresis	99.0%	99.96%	99.9%	100%

Table 1. *Worldwide availability results.*

LIMITATIONS OF THIS ANALYSIS

Correlation

The current results have assumed that there is no temporal correlation between pseudorange errors after two minutes. This is a rather optimistic assumption as the clock and ephemeris errors have a much longer correlation time (on the order of hours [4]). This fact does not undermine the current results. Even though the pseudorange errors might be correlated in time it does not necessarily mean that the test statistics are, since the geometry changes substantially in the course of an hour. Also, even if clock and ephemeris errors are very correlated in time, the URA (which represents an overbound) is very conservative for the core of the distribution, which is what matters most in this analysis. This being said, it will be necessary to develop an approach to evaluate in a conservative way the loss due to correlation.

Nominal biases

The nominal biases that were assumed in [1] have not been included in this analysis. The inclusion of the biases will degrade both the baseline and the hysteresis results.

Algorithm complexity

As it has been presented, the algorithm is computationally intensive. For each satellite there are 30 possible branches in the threat model. Since there are on average 10 satellites in view, this means that there 300 terms in the sum (21). Each term in this sum requires a worst case bias *b*. For each evaluation of an error bound in the VPL search there are 300 bias searches. One approach to simplify the algorithm is to find approximations that simplify the worse bias search. It is also possible that most of the benefit in this algorithm comes from the geometry change and not from the simple accumulation of test statistics. This would greatly simplify the algorithm because we would only need to find the best test for each of the fault tree branches.

CONCLUSION

This work has presented a conservative threat model for failures in the time domain which is a generalization of ARAIM threat model. A variant of the ARAIM algorithm was developed to account for this temporal threat model in the Protection Level calculation. With this algorithm we have evaluated the performance gain for several single constellations in the single failure case, and we have found that accounting for hysteresis could significantly improve the coverage of LPV 200 – in some important cases the unavailability is reduced by 50%. Because of this, the use of hysteresis in the VPL calculation can be one more tool to increase ARAIM performance and acceptance for vertical guidance. .

Several assumptions presented here are only a starting point for further discussion and analysis. For example, the present study has ignored the correlation that exists between the test statistics taken at different epochs, so further work will be necessary to evaluate reduction in VPL in the presence of correlation. As a consequence the magnitude of the improvement presented here can be interpreted as an upper bound. It is also important to note that the VPL computation will need to be simplified to be viable in an airborne receiver.

ACKNOWLEDGEMENTS

This work was sponsored by the FAA GPS Satellite Product Team (AND-730).

APPENDIX

Independence of test statistic noise to position error noise

The vectors of coefficient s and s_{subset} are defined as:

$$s = u^T \left(G^T W G\right)^{-1} G^T W \tag{A.1}$$

$$s_{subset} = u^T \left(G^T W_{subset} G\right)^{-1} G^T W_{subset} \tag{A.2}$$

W_{subset} is the same as W for all entries except for the diagonal term corresponding to the excluded satellite, which is zero. For the definitions of the variables, see [1]. We have:

$$\operatorname{cov}\left(\left(s - s_{subset}\right)^T \varepsilon, s^T \varepsilon\right) = \left(s - s_{subset}\right)^T \operatorname{cov}\left(\varepsilon\right) s \tag{A.3}$$

Since we have:

$$\operatorname{cov}\left(\varepsilon\right) = W^{-1} \tag{A.4}$$

We have:

$$\begin{aligned} &\operatorname{cov}\left(\left(s - s_{subset}\right)^T \varepsilon, s^T \varepsilon\right) = \\ &u^T \left(\left(G^T W G\right)^{-1} G^T W - \left(G^T W_{subset} G\right)^{-1} G^T W_{subset}\right) \\ &W^{-1} W G^T \left(G^T W G\right)^{-1} u \\ &= u^T \left(I_4 - \left(G^T W_{subset} G\right)^{-1} G^T W_{subset} G^T\right)\left(G^T W G\right)^{-1} u = 0 \end{aligned} \tag{A.5}$$

REFERENCES

[1] T. Walter et al. “Worldwide Vertical Guidance of Aircraft Based on Modernized GPS and New Integrity Augmentations” to appear in Proceedings of the IEEE Special Issue.

[2] J. Blanch, et al. “RAIM with Optimal Integrity and Continuity Allocations Under Multiple Satellite Failures,” submitted to IEEE Transactions on Aerospace Electronics.

[3] S. S. Jan et al. “Matlab Simulation Toolset for SBAS Availability Analysis,” in Proceedings of the ION GPS-01, Salt Lake City, 2001.

[4] T. Walter et al. “Evaluation of Signal in Space Error Bounds to Support Aviation Integrity,” in Proceedings of the ION GNSS 09, Savannah, 2009.