Designing integrity into position estimation

P.B. Ober
Delft University of Technology - Telecommunications and Traffic Control Systems Group
Mekelweg 4, 2628 CD Delft, The Netherlands

ABSTRACT

While position *integrity* is crucial when positioning systems are to be used for safety critical operations such as in aviation applications, currently known positioning algorithms are generally optimised for *accuracy* instead. Even when they are combined with fault detection and exclusion schemes, these algorithms still give sub-optimal integrity. The paper therefore promotes a new way of algorithm design that takes integrity rather than accuracy as the parameter to optimise.

The new class of high integrity positioning algorithms that is thus described aims at obtaining improved integrity with both current and new systems; not by improving the physical infrastructure, but by using clever algorithmic optimisation in the receiver. A small simulation example shows that the integrity and availability of unaugmented GPS for non-precision approach can indeed be improved substantially.

1. INTRODUCTION

This paper will discuss integrity of navigation systems, focussing on general methods to design integrity into a navigation system. This is of particular importance for safety critical applications such as aviation. Navigation literature still lacks a systematic overview of general ways to obtain a certain amount of integrity in a system. Many books and papers [Sturza88][Brenner90][Brown96][Leva96] analyse given fault detection and exclusion (FDE) schemes without considering the existence of alternatives. This paper therefore aims at broadening the current views by describing a systematic and general approach to design integrity into a system's algorithms.

The paper first introduces the navigation performance parameters including integrity parameter, and shows that its meaning has been changed over the years. It then discusses the algorithms in positioning systems and shows that integrity is a function of two algorithms: the position computation and the error detector. Current implementations of these algorithms optimise for accuracy, which leads to sub-optimal integrity. Therefore, the paper advocates a different design strategy that uses integrity as a starting point.

2. INTEGRITY IN THE REQUIRED NAVIGATION PERFORMANCE CONCEPT

Required Navigation Performance (RNP) is a concept designed to achieve a desired Target Level of Safety (TLS). It describes the joint performance of the navigation sensor and the flight control system (FCS). As a result, it allows for different ways to achieve the required performance, as the performance requirements can be distributed freely between the navigation sensor and the FCS.

2.1 The original concept

The original concept from [Kelly94] starts with defining the outer tunnel, a containment surface in space, centred on the assigned flight path that defines the obstacle clearance, terrain avoidance or aircraft separation criteria. The aircraft should remain within the outer tunnel with a probability near to one. When the aircraft leaves the outer tunnel unintentionally, this is called a tunnel incident. Since the events that cause a tunnel incident are uncertain, the tunnel incident must be quantified by probabilities called risks. The concept also defines an inner tunnel that relates to the nominal system performance and should contain the aircraft most (typically 95%) of the time.

RNP defines a total of four performance parameters: accuracy, integrity, continuity and availability, that each corresponds to the risk of a certain event that could cause a tunnel incident. Accuracy covers the risk that excessive system error causes a tunnel incident. The risk associated with latent system failures is covered by the integrity requirement, while the risk of an unscheduled guidance function loss is specified by the continuity requirement. Finally, availability covers the risk of a lack of guidance at the start of the operation. The definition of accuracy reads as follows:

> *Accuracy* is the ability of the total system to maintain the aircraft position *within the inner and outer tunnel* with a sufficiently high probability.

[Kelly94] has allocated probabilities of 0.95 for the inner tunnel, and of $1\text{-}10^{-7}$ per approach for the outer tunnel. The 'ability' in the accuracy requirement refers to a correctly working system only. Note that the requirement to stay within the inner tunnel is not directly related to the tunnel incident risk and is therefore of a rather different nature than the other RNP requirements.

Integrity is defined as follows:

> *Integrity* is that quality which relates to the trust that can be placed in the correctness of the information supplied by the total system. Integrity risk is the probability of an undetected (latent) failure of the specified accuracy. Integrity includes the ability of the system to provide timely warnings to the user when the system should not be used for the intended operation. Such a warning is called a tunnel incident alarm.

As is clear from the definition, integrity is related to the system's capability to generate alarms in situations of insufficient accuracy. This capability might vary with time and should therefore be monitored by an integrity monitoring system. When the integrity risk becomes too high, the user should be notified that there is insufficient guarantee that a timely alarm can be generated in case of lacking accuracy.

2.2 Changes in the definitions

Since the publication of the original RNP concept other definitions of integrity and accuracy have been introduced [DO-236][DO-229]. In [DO-236] the accuracy requirement is given as:

> Each aircraft operating in RNP airspace shall have total system error components in the cross-track and along track directions that are less than the RNP value 95% of the flying time.

while the integrity requirement reads:

> The probability that the total system error of each aircraft operating in RNP RNAV airspace exceeds the specified cross-track containment limit without annunciation shall be less than 10^{-5} per flight hour. The cross track containment limit is twice the RNP RNAV value.

These definitions will be used in the remainder of the paper. Note that accuracy now refers *only* to the nominal system performance and is no longer specified at the containment limit level. Furthermore, integrity risk is no longer the probability of an undetected lack of accuracy, but refers to the probability of an undetected violation of the containment region instead. As a result, the integrity alarm mechanism no longer seems to apply to the nominal system performance. It seems that a lack of nominal performance is no longer considered a safety issue, as long as the probability of announced violation of the containment limit remains sufficiently small.

We can conclude that integrity has become even more important than it was in the original RNP concept, as it now also includes the probability of a tunnel incident when the system works properly. Therefore, optimising for integrity instead of accuracy now immediately implies minimising the probability of tunnel incidents.

3. SYSTEM DESIGN PARAMETERS

As a natural consequence of the way performance is quantified, a position system's design goals can be defined in terms of achieving the best possible accuracy, integrity, continuity and availability. All parameters will benefit from the addition of more position information in the form of more or better signals, better measurement geometry or a removal of certain failure modes. Therefore, adding more position information is the obvious, but sometimes prohibitively expensive, solution to improve performance and is beyond the scope of this paper, that deals exclusively deals with the design of optimal *algorithms* within a system with a given number of signals of a certain quality, geometry and failure rate.

In general, one could say that within the constraint of such a given infrastructure - there is a trade-off between all four RNP parameters. A realistic system design should therefore optimise for one parameter within minimum requirement constraints for the others. Up till now, the usual way to design positioning algorithms has been to optimise *accuracy*, while all other parameters are evaluated *a posteriori* to see whether they obey the requirements. As we will have shown, integrity would be the parameter of choice to optimise systems for safety critical operations. The following sections will therefore investigate how such an optimising might be obtained.

4. POSITIONING SYSTEM ALGORITHMS

In general, integrated navigation systems aggregate data from a variety of sensors and databases to estimate position or position-related parameters like differential corrections or measurement biases. While the RNP parameters relate to position only, they can (and are) also be used in a general parameter estimation context. Therefore, everything that will be said here on position estimation can be applied to the estimation of other parameters as well.

In the most general set-up, a positioning system contains four different functions that together should provide a parameter estimate with sufficient accuracy, integrity and continuity:

1. A positioning algorithm
2. An error detection algorithm
3. An error identification algorithm
4. An integrity monitor

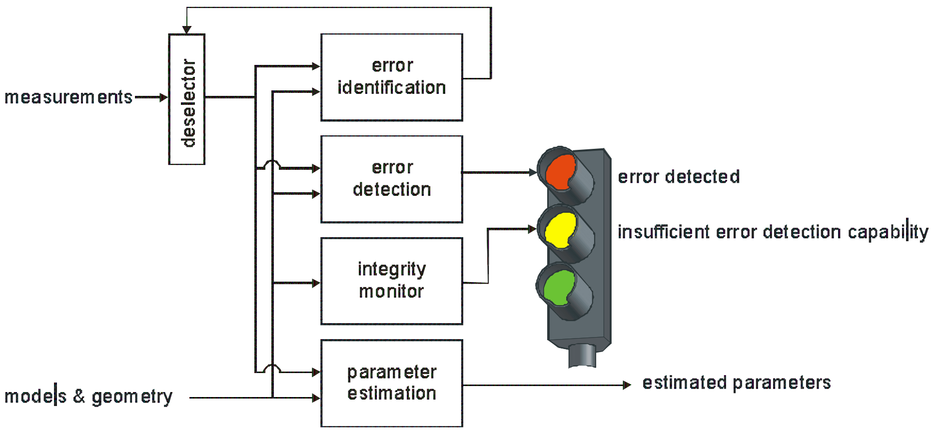

Figure 1. The traffic light analogy of a navigation system and the relation between its algorithms and its outputs / states

Error detection attempts to provide an alarm whenever there is a position failure. It is usually combined with error identification, that tries to identify the erroneous measurements and removes them from the position computations. The integrity monitor is the algorithm that determines the performance of the error detection algorithm, and warns the pilot in case of a lack of error detectability. To understand the different system states that can exist, it is illustrative to compare the system to a traffic signal, see Figure 1. While a green light indicates the presence of sufficient integrity, "yellow" indicates a lack of integrity due to insufficient failure detection power. The red light corresponds to the tunnel incident alarm and thus stands for a detected failure. Note that the yellow light already means, that the system can not be guaranteed to be safe, which is -in a way- true for ordinary traffic signals as well.

The required performances of the positioning, error detection and error identification algorithms are strongly related to each other. Because error identification is meant for improving continuity, we will only focus on positioning and error detection. Obviously, the better the positioning algorithm can resist measurement failures, the lower the amount of position failures and the less error detection capability will be required. In the case that the position failure rate drops below the integrity requirements, the need for an error detection algorithm effectively disappears. This shows that the integrity requirement can in fact be distributed among these two algorithms.

This same point is visualised in Figure 2. As the positioning algorithm influences the position error, it affects both accuracy and integrity, while the error detection algorithm affects the error detection signal and therefore integrity. As integrity is concerned with detection of position errors rather than measurement errors, *integrity is in fact determined by the relationship between position error and error detection signal*. The “stronger” this connection is the better integrity gets. Current positioning systems however tend to focus on the links between measurement error and position, and measurement error and error detection signal separately, and optimise for accuracy rather than integrity, as we will see in the next section.

4.1 Current approach

Current positioning algorithms are almost always least squares estimation schemes. When the system is linear and measurement noise is normally distributed, this gives optimal accuracy: it is the best way to mitigate the effects of noise on the position. On the other hand, the obtained position is sensitive to failures: a single wrong measurement can cause an arbitrarily large position error. To improve on this, error detection has to be used.

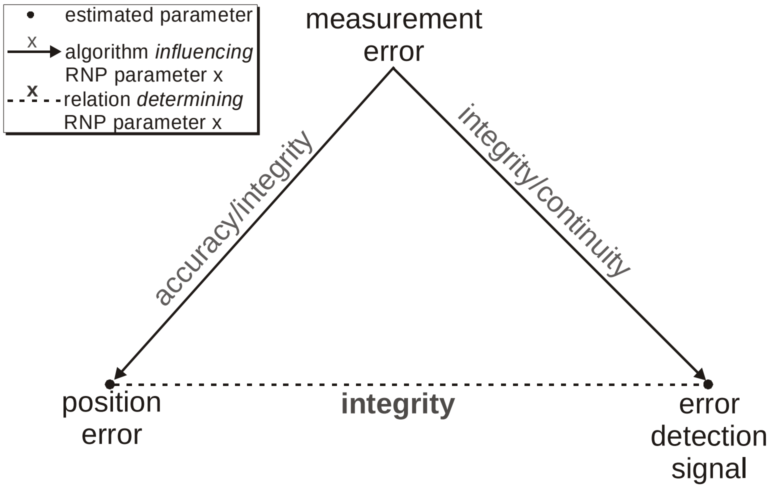

Figure 2. The relations between RNP performance parameters, measurement error, position error and error detection.

For error detection usually a test statistic is used that is based on the least squares residual; this residual is the most accurate estimate of the measurement error. When the residual becomes too large, an error is detected. It can be proven that the least squares residual and the position error are statistically independent. The reason that detection still works is that both residual and position are influenced by the same –deterministically modelled- bias. However, the noise in the position error is not reflected at all in the residual, indicating that the error detection properties – and therefore integrity- might not be optimal.

4.2 New approach

In order to optimise integrity, a new approach to algorithm design is required that focuses on the connection between position error and error detection signal, rather than on positioning and error detection separately. We propose to refer to this kind of algorithms as high integrity positioning (HIP) [Ober99]. Although the HIP framework is only in the starting phase of its development, a prototype algorithm has been implemented that searches for a position solution that has optimal integrity while being sufficiently accurate. In other words: integrity – measured by virtue of the Horizontal Protection Level [Leva96] is optimised under the explicit condition that a sufficiently low horizontal dilution of precision (HDOP) is maintained. A small simulation has been performed to obtain an impression of the performance improvements that might be obtained.

In the simulations, non-augmented GPS performance for lateral navigation has been compared to the requirements for non-precision approach. The simulation is performed with a nominal 24-satellite GPS configuration and the GPS and RNP parameters from [DO-208]. GPS performance parameters are computed over a world-wide 3-degree grid, giving a total of about 6500 data points. Table 1 summarises the mean performance parameters of both HIP and the traditional least squares. The advantages of HIP are clearly expressed in the largely improved system availability that is obtained thanks to the higher integrity. Note that there is –as was to be expected- a small loss of accuracy. This loss, however, is controlled by the algorithm, and never leads to unavailability.

Table 1. Overall comparison of HIP and traditional least squares

Performance parameter	**HIP**	**Least Squares**
Unavailability	0.027	0.13
Integrity (mean HPL in meters)	245	356
Accuracy (mean HDOP)	1.6	1.3

5. CONCLUSIONS AND RECOMMENDATIONS

We have seen that using the current interpretation of RNP parameters, integrity has become the prime safety parameter. Current navigation algorithms are optimised to give optimal accuracy, while integrity is only evaluated afterwards. Therefore, when one wants to optimise integrity, a new way of thinking is required, focussing on the connection between position error and error detection signal rather than on positioning and error detection separately. This way, it becomes possible to develop methods to design integrity into a system's algorithms and achieve High Integrity Positioning (HIP).

First results show that a substantial integrity improvement is possible. The first developed HIP algorithm optimises integrity under the explicit condition that sufficient accuracy is maintained, and reduces the non-availability of standalone GPS for non-precision approach from 13% to 2.7% when compared to the accuracy optimised least squares algorithm. The integrity, measured by the protection limit, improved by 50%. Possibly, these figures can still be improved upon when HIP becomes a more mature technique, or when the accuracy requirements –that have no direct safety impact in the current RNP context- are relaxed to give the algorithms more freedom to optimise integrity.

To conclude, we express the hope that the new way of thinking promoted in this paper might assist in exploring a whole new class of algorithms, obtaining improved integrity with both current and new systems; not by improving the physical infrastructure, but by using clever, integrity optimised positioning and error detection algorithms.

REFERENCES

[Brenner90] Brenner, M; "Implementation of a RAIM monitor in a GPS Receiver and an Integrated GPS", Proceedings of the Third International Satellite Division Meeting, Institute of Navigation, September 1990

[Brown96] Brown, R.G.; "Receiver Autonomous Integrity Monitoring", in *Global Positioning System: theory and applications*, Volume 2, Edited by B.W. Parkinson and J.J. Spilker Jr., American Institute of Aeronautics and Astronautics, Inc., Washington DC, 1996

[DO-208] Minimum Operational Performance Standards for Airborne Supplemental Navigation Equipment Using Global Positioning System (GPS), RTCA Inc., July 1991, USA

[DO-229] RTCA DO-229 DO-229A, Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment

[DO-236] RTCA DO-236, Minimum Aviation System Performance Standards: Required Navigation Performance for Area Navigation, issued 1-27-97

[Kelly94] Kelly, R.J. and J.M. Davis; "Required Navigation Performance (RNP) for Precision Approach and Landing with GNSS Application", Journal of the Institute of Navigation, Vol. 41, No. 1, pp. 1-30, 1994

[Leva96] Leva, J.L., M. Uijt de Haag and K. Van Dyke; "Performance of Standalone GPS", in Understanding GPS: Principles and Applications, Edited by E.D. Kaplan, Artech House, 1996

[Ober99] Ober, P.B. "Towards High Integrity Positioning", Proceedings of the 12th International Technical Meeting of The Satellite Division of The Institute of Navigation, ION GPS-99, Nashville TN, USA

[Sturza88] Sturza, M.A.; "Navigation System Integrity Monitoring using redundant measurements", Navigation, Vol. 35, No. 4, Winter 1988-89