

The Discovery and Quantification of Risk in High Dimensional Search Spaces

Kester Clegg & Rob Alexander

University of York

York, UK

{kester.clegg | rob.alexander}

@york.ac.uk

ABSTRACT

We describe a technique used by the ASHiCS project (Automating the Search for Hazards in Complex Systems) to discover high risk air traffic control (ATC) scenarios. We use a fast time ATC simulation of an en-route sector containing multiple flight paths and aircraft types, and into this we inject a serious incident (cabin pressure loss) which forces one aircraft to make an emergency descent. In order to create additional workload for the air traffic controller (ATCo), we also introduce a storm moving across the sector. We measure the associated levels of risk by analyzing the simulation outputs, selecting scenarios on basis of most risk and mutating aircraft entry times to see if the search can discover variant scenarios of even greater risk. The search space is extremely large and cannot be exhaustively searched for the worst case; this is a problem for safety engineers who require a context to search results so that event probabilities can be determined. While providing context cannot demonstrate that the worst case scenario has been found over all input permutations, it can indicate the expected frequency of that result in its near neighborhood, allowing analysts to focus on a much reduced parameter range when investigating those aircraft in conflict.

Categories and Subject Descriptors

I.2.8 [ARTIFICIAL INTELLIGENCE]: Problem Solving, Control Methods, and Search – *Heuristic methods*

I.6.3 [SIMULATION AND MODELING]: Applications

General Terms

Algorithms; Management; Reliability; Security

Keywords

Air traffic control/management; safety; heuristic search; simulation

1. INTRODUCTION

With increasingly complex systems to manage, safety engineers are increasingly concerned that their analysis tools are not adequate for them. ASHiCS was set up to develop a proof-of-concept tool that uses evolutionary search and a fast-time air traffic control (ATC) simulation to uncover hazards that might otherwise be missed using traditional manual safety analysis.

We use a fast time ATC simulator (RAMS Plus – <http://www.ramsplus.com/>) configured with a model of an en-route sector containing multiple flight paths and aircraft types. Into this we inject a serious incident (cabin pressure loss), which requires one aircraft to make an emergency descent. To create additional workload for the air traffic controller (ATCo), we also introduce a storm moving across the sector. We measure the associated levels of risk by analyzing the simulation outputs, selecting scenarios on basis of most risk and mutating aircraft entry times to see if the search can discover scenarios of still greater risk.

2. SEARCHING FOR RISK

The ASHiCS search harness uses a population of 50 scenarios per generation and runs for up to 250 generations. A scenario is defined by assigning flightpaths and sector entry times to each of 20 aircraft. The search implements what is sometimes termed a “near neighbor, random hill-climber” search. That is, individual scenarios are selected and the aircraft start times are mutated within a limited range that defines the size of the “neighborhood” of the result (generally a few minutes either side of the previous entry time). Provided such mutations are not too radical we should be guaranteed that a “near neighbor” of the original scenario is created, as aircraft remain on their flight paths, relatively close to their previous start times. However, to ensure that the search has not been unlucky in its initial seeding of random samples (in what is a very large configuration space), we continue to allow a proportion of each population to be generated entirely by random sampling. The split in the population is dictated by a policy of elitism; the top twenty per cent of a generation’s scenarios are each mutated to create three “near neighbors” which are carried over to the next generation. The remaining forty per cent of the population is created from new random samples.

A mid-air collision is the worst possible outcome from a run, but this is extremely unlikely. Rather than looking for collisions, therefore, the search attempts to maximize a fitness function composed from a range of factors that increase the risk of a collision:

- total number of conflicts (instances where two aircraft flight paths risk crossing or coming too close);
- closest approach distance of any two aircraft (as a percentage of minimum safe separation);
- total number of resolutions by an ATCo;
- total ATCo task workload (measured in seconds);
- the NASA sector complexity measure for the airspace (highest value observed for this measure during the run)

Given this function and the search method above, it was not difficult to find situations that lead to high risk measures. Figure 1 illustrates a typical progression over the course of a 1200-individual search – the highest risk score so far (shown by the horizontal lines) steadily climbs as the search progresses.

Copyright is held by the author/owner(s).

GECCO’13 Companion, July 6–10, 2013, Amsterdam, The Netherlands.

ACM 978-1-4503-1964-5/13/07.

3. FINDING THE WORST CASE

Finding a high-risk situation that wasn't previously known is indeed interesting, and could be useful to ATM system designers. Designers will want to know more, however: is this the worst case among all similar cases? Or are there many more variants to discover? Exhaustive search could answer these questions, but high dimensional solution spaces are very expensive to search exhaustively, and doing so would in any case make our heuristic search redundant.

We therefore looked at how we could gain a better understanding of the context of our result by conducting sensitivity analysis of the best scenarios within the evolving population [1]. We discovered that the search performance was linked to the mutation range operator, suggesting it was sensitive to radical mutations as these performed less well. The picture this presented was one of relatively few high scoring scenarios that formed "peaks" with steep, narrow sides. The only way to climb such peaks was through small incremental mutations. However we still didn't know how many solutions lay in the near neighbourhood, or whether we what we had found was close to the worst case scenario. As this type of information is crucial to safety analysts, we decided to extend the idea of a sensitivity analysis by applying it to the area around the result of the search.

Thus, the second stage of our process extensively samples the near neighborhood of the final result using random configurations (random within some defined tolerance of the first stage highest-risk configuration). As we are able to sample many thousands of scenarios within this small parameter range, we are able to achieve coverage levels far in excess of those obtained during the original search and so gain an idea of how many other high scoring scenarios lie in the near neighbourhood of the original result.

From Fig. 1 and 2 we can see that the search (the first stage) failed to find the worst case scenario. By extensively sampling the near neighborhood (as the second stage), we were able to uncover 3 variants whose aircraft entry times improved the fitness score of the original scenario from Stage 1 – these are the points above the horizontal line in Figure 2. However, it is also apparent that the original search result is very close to these worse cases. It would appear that although some marginal improvement is possible, the search performed well in terms of finding the worst case scenario.

The second stage provides some insight to the nature of the solution space within the near neighborhood of the original result in terms of the frequency of high risk scenarios and how those scenarios differ from the original result. By analyzing the high risk variants, safety analysts can focus on the parameter range that generates the worst cases and can then investigate how to prevent that configuration of inputs leading to a hazard in the air sector being modeled.

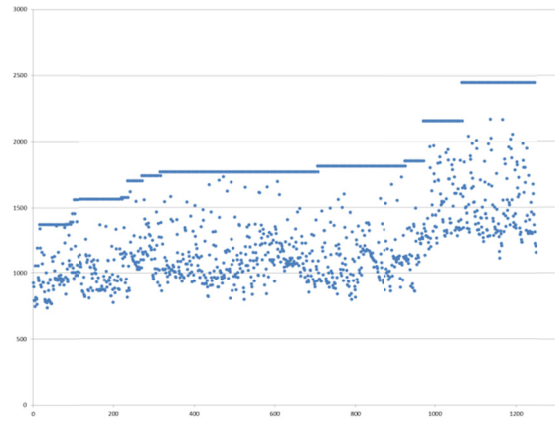


Figure 1: Stage 1 sensitivity analysis for storm scenarios. Mutation range = 30s.

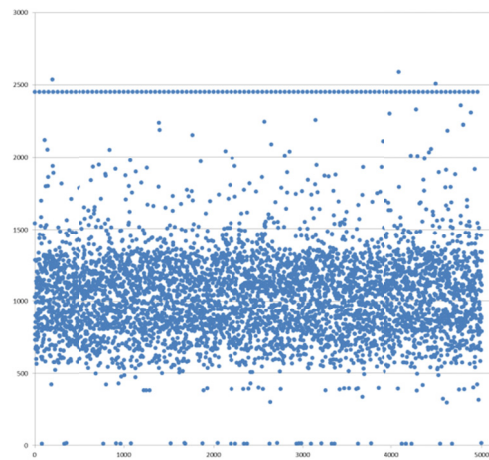


Figure 2: Stage 2 near neighbor sampling of original search result.

4. FINDING ALL HAZARDS

System designers and safety engineers may well ask a harder question: how do we know when the search has found all the dangerous situations in the search space? This remains a problem, and is a key topic for our further research.

5. ACKNOWLEDGEMENTS

This work was co-financed by EUROCONTROL acting on behalf of the SESAR Joint Undertaking (the SJU) and the EUROPEAN UNION as part of Work Package E in the SESAR Programme. Opinions expressed in this work reflect the authors' views only, and EUROCONTROL and/or the SJU shall not be considered liable for them or for any use that may be made of the information contained herein.

6. REFERENCES

- [1] Clegg, K., Alexander, R. Searching for Risk in Large Complex Spaces. In *Proceedings of EvoStar*, 2013