Estimating Probability of Failure of a Complex System Based on Inexact Information about Subsystems and Components, with Potential Applications to Aircraft Maintenance

V. Kreinovich[1], C. Jacob[2,3], D. Dubois[3]
J. Cardoso[3], M. Ceberio[1], and I. Batyrshin[4]

[1]University of Texas, El Paso, Texas, USA
[2]Institut Supérieur de l'Aéronautique et de l'Espace (ISAE), Toulouse, France
[3]Institute de Recherche en Informatique de Toulouse (IRIT), France
[4]Instituto Mexicano de Petróleo, México D.F.

emails vladik@utep.edu, jacob.christelle.11@gmail.com, {cardoso,dubois}@irit.fr, mceberio@utep.edu, batyr@imp.mx

1. Formulation of the Practical Problem

- In practice, while it is desirable to minimize risk, it is not possible to completely eliminate it.
- There are always some very low probability events that can potentially lead to a system's failure.
- All we can do is to make sure that the probability of failure does not exceed the desired small value p_0.
- For example, the probability of a catastrophic event is usually required to be at or below $p_0 = 10^{-9}$.
- As a result, we need to estimate the probability of failure p_f of a complex system.
- E.g., we need to estimate p_f of an aircraft:
  - on design stage, to check that the design is safe;
  - on working stage, to check whether maintenance is needed (i.e., whether $p_f \le p_0$).

2. Information Available for Estimating System's Probability of Failure

- Complex systems consist of subsystems, which, in turn, consist of components (or maybe of sub-subsystems).
- So, to estimate the probability of failure of a complex system, we need to take into account:
  - when the failure of components and subsystems lead to the failure of the system as a whole, and
  - how reliable are these components and subsystems.
- Complex systems are usually built with redundancy.
- Usually, we know the structure of the system.
- As a result, for each possible set of failed components, we can tell whether this set will lead to a system failure.
- In this paper, we will assume that this information is available.

3. How Reliable Are Components and Subsystems?

- For each component A, there is a probability $P(A)$ of its failure.
- In general, we can estimate $P(A)$ by testing.
- Specifically, $P(A)$ can be estimated as the frequency of cases when the component failed.
- Aircrafts are usually built of highly reliable components.
- Hence, only a few failure cases are available, not enough for an accurate estimate of $P(A)$.
- So, we only have a confidence interval $\mathbf{P}(A) = [\underline{P}(A), \overline{P}(A)]$ containing the actual (unknown) value $P(A)$.

4. Component Failures Are Independent Events

- In many practical situations, failures of different components are caused by independent factors.
- For example, for an aircraft:
  - possible failures of mechanical subsystems can be caused by the material fatigue, while
  - possible failures of electronic systems can be caused by atmospheric electricity (e.g., a thunderstorm).
- In this paper, we assume that failures of different components are independent events.
- Our objective is:
  - to use all this information
  - for estimating the probability of failure of the entire complex system.

5. Simplest Case: Component Failures are Independent, Failure Probabilities $P(A)$ Are Known

- *Main idea:* Monte-Carlo simulations.
- For each component A, we simulate a Boolean variable *failing(A)* which is:
  - true with probability $P(A)$ and
  - false with the remaining probability $1 - P(A)$.
- Then, we use the known redundancy structure to check whether the system as a whole fails or not.
- If the system failed in f out of N simulations, we estimate p_f as f/n.
- From statistics, we can determine N.
- *Example:* to estimate p_f with accuracy 10% and certainty 99.9%, we need $N \geq 225$.

6. Important Subcase of the Simplest Case: When Components are Very Reliable

- In aircrafts, components are highly reliable: $P(A) \ll 1$.
- As a result, with high probability, in all simulations, all the components will be simulated as working properly.
- Thus, to estimate p_f as f/N, we need a large number of iterations N.
- To speed up simulations, instead of $P(A)$, we use re-scaled (larger) values $\lambda \cdot P(A)$ for some $\lambda \gg 1$.
- The probability $p_f(\lambda)$ depends on λ as $p_f(\lambda) \approx \lambda^k \cdot p_f$ for some positive integer k; so, we:
  - repeat this procedure for two different values $\lambda_1 \neq \lambda_2$, get the two values $p_f(\lambda_1)$ and $p_f(\lambda_2)$, and
  - find both unknown k and p_f from the resulting system of two equations with two unknowns: $p_f(\lambda_1) \approx \lambda_1^k \cdot P$ and $p_f(\lambda_2) \approx \lambda_2^k \cdot P$.

7. Interval values $[\underline{P}(A), \overline{P}(A)]$: Monotonicity Case

- In many situations:
  - if for a certain list of failed components, the system fails,
  - then it will still fail if we add one more components to the list of failed ones.
- In this case, the smaller the probability of failure $P(A)$ for each component A, the smaller the probability p_f.
- Thus:
  - to compute $\underline{p}_f$, we assume that for each component A, the failure probability is equal to $\underline{P}(A)$;
  - to compute $\overline{p}_f$, we assume that for each component A, the failure probability is equal to $\overline{P}(A)$.

8. In Practice, the Dependence is Sometimes Non-Monotonic

- Often, the dependence of the system's failure on the failure of components is non-monotonic:
  - adding one more failing component to the list of failed ones
  - suddenly makes the previously failing system recover.
- This non-monotonicity is caused by redundancy: there is a back-up system.
- When only one important sensor fails, the system as a whole may fail.
- However, if many sensors fail, the system detects failure and invokes a back-up system.
- In this paper, we will consider the non-monotonic case, when the above simple algorithm is not applicable.

9. When Dependence Is Non-Monotonic: General Analysis

- For each component A, the formula of full probability implies $p_f = P(A) \cdot P(F|A) + (1 - P(A)) \cdot P(F|\neg A)$.
- The resulting dependence of P on $P(A)$ is linear.
- A linear function attains it minimum and maximum at the endpoints.
- Thus, to find $\underline{p}_f$ and $\overline{p}_f$, it is sufficient to only consider two values for each A: $P(A) = \underline{P}(A)$, $P(A) = \overline{P}(A)$.
- For k components, we need 2^k combinations of such endpoints; for large k, this is not feasible.
- In general, the problem of computing the range $[\underline{p}_f, \overline{p}_f]$ is proven to be NP-hard.
- This means, crudely speaking, that the exponential time is inevitable.

10. A Practically Important Case of Narrow Intervals

- *Usually:* the intervals $[\underline{P}(A), \overline{P}(A)]$ are narrow.
- *Notations:* $\widetilde{P}(A)$ is the interval's midpoint, and $\Delta(A)$ its radius.
- *Narrow means:* terms quadratic (and of higher order) in $\Delta P(A) \stackrel{\text{def}}{=} P(A) - \widetilde{P}(A)$ can be safely ignored.
- *Idea:* expand $p_f = p_f(\widetilde{P}(A) + \Delta P(A), \ldots)$ into Taylor series and keep only linear terms:

$$p_f \approx \widetilde{p}_f + \sum_A c_A \cdot \Delta P(A), \text{ where } c_A \stackrel{\text{def}}{=} \frac{\partial p_f}{\partial P(A)}.$$

- *Conclusion:* $\Delta p_f \stackrel{\text{def}}{=} p_f - \widetilde{p}_f \in [\Delta, \Delta]$, where

$$\Delta \stackrel{\text{def}}{=} \sum_A |c_A| \cdot \Delta(A).$$

11. Case of Narrow Intervals (cont-d)

- *Natural idea:* compute all the partial derivatives c_A and to use the above formula for Δ.
- *Example:* pick small h and use numerical differentiation

$$\frac{\partial p_f}{\partial P(A)} \approx \frac{p_f(P(A)+h, P(B), \ldots) - p_f(P(A), P(B), \ldots)}{h}.$$

- *Problem:*
  - each computation of p_f requires time-consuming Monte-Carlo simulations;
  - we need to repeat this computation as many times as there are components;
  - an aircraft has thousands of components, so computations are too long.

12. Cauchy Deviate Techniques: Reminder

- *Cauchy distribution with parameter* Δ *(reminder):* probability density function $\rho(z) = \dfrac{\Delta}{\pi \cdot (z^2 + \Delta^2)}$.
- *Property:*
  - if z_A are independent Cauchy random variables with parameters $\Delta(A)$,
  - then $z = \sum\limits_A c_A \cdot z_A$ is Cauchy with parameter

$$\Delta = \sum_A |c_A| \cdot \Delta(A).$$

- Therefore, using Cauchy distributed random variables δ_A with parameters $\Delta(A)$, the difference

$$c \stackrel{\text{def}}{=} p_f(\widetilde{P}(A) + \delta_A, \ldots) - p_f(\widetilde{P}(A), \ldots) = \sum_A c_A \cdot \delta_A$$

is Cauchy distributed with the desired parameter Δ.

13. Cauchy Deviate Techniques (cont-d)

- Using Cauchy distributed random variables δ_A with parameters $\Delta(A)$, the difference

$$c \stackrel{\text{def}}{=} p_f(\widetilde{P}(A) + \delta_A, \ldots) - p_f(\widetilde{P}(A), \ldots) = \sum_A c_A \cdot \delta_A$$

  is Cauchy distributed with the desired parameter Δ.

- So:
  - repeating this experiment N_c times, we get N_c values $c^{(1)}, \ldots, c^{(N_c)}$;
  - these values are Cauchy distributed with the unknown parameter Δ;
  - from these values, we can estimate Δ.

14. Cauchy Deviate Techniques: Important Comment

- *To avoid confusion:* we should emphasize that the use of Cauchy distributions is:
  - a computational technique,
  - *not* an assumption about the actual distribution.
- *Explanation:*
  - we know that the actual value of $\Delta P(A)$ is bounded by $\Delta(A)$;
  - however, for a Cauchy distribution, there is a positive probability that the simulated value is $> \Delta(A)$.

15. Cauchy Techniques: Towards Implementation

- Simulating Cauchy distr.: $\delta_A = \Delta(A) \cdot \tan(\pi \cdot (r_A - 0.5))$, where r_A is uniform on $[0, 1]$.
- In order to estimate Δ, we apply the Maximum Likelihood Method $\rho(c^{(1)}) \cdot \rho(c^{(2)}) \cdot \ldots \cdot \rho(c^{(N_c)}) \to \max$.
- Equating derivative to 0, we get:

$$\frac{1}{1 + \left(\frac{c^{(1)}}{\Delta}\right)^2} + \ldots + \frac{1}{1 + \left(\frac{c^{(N_c)}}{\Delta}\right)^2} = \frac{N_c}{2}.$$

- The left-hand side is increasing,
  - it is equal to 0 (hence $< N_c/2$) for $\Delta = 0$, and
  - it is $> N_c/2$ for $\Delta = \max \left|c^{(k)}\right|$.
- Therefore the solution to this equation can be found by applying a bisection method to the interval $\left[0, \max \left|c^{(k)}\right|\right]$.

16. Implementation: Details

- We assumed that the function p_f is reasonably linear when the values δ_A are small: $|\delta_A| \leq \Delta(A)$.
- However, the simulated values δ_A may be larger than $\Delta(A)$.
- When we get such values, we do not use the original function p_f.
- For such values, we use a "normalized" function:
  - that is equal to p_f within the given intervals, and
  - that is extended linearly for all other values.

17. Cauchy Deviate Technique: Main Algorithm

- Apply p_f to the values $\widetilde{P}(A)$ and compute $\widetilde{p}_f = p_f(\widetilde{P}(A), \ldots)$.
- For $k = 1, 2, \ldots, N_c$, repeat the computation of the following values:
  - n random numbers $r_A^{(k)}$ uniformly distrib. on $[0, 1]$;
  - Cauchy distributed values $c_A^{(k)} = \tan(\pi \cdot (r_A^{(k)} - 0.5))$;
  - the largest $K = \max\limits_A |c_A^{(k)}|$;
  - the simulated deviations $\delta_A^{(k)} := \Delta(A) \cdot c_A^{(k)}/K$;
  - simulated probabilities $P^{(k)}(A) = \widetilde{P}(A) + \delta_A^{(k)}$;
  - $p_f(P^{(k)}(A), P^{(k)}(B), \ldots)$ and

$$c^{(k)} = K \cdot (p_f(P^{(k)}(A), P^{(k)}(B), \ldots) - \widetilde{p}_f);$$

- Finally, compute Δ by applying the bisection method to solve the corresponding equation.

18. Resulting Gain and Remaining Limitations

- The Cauchy deviates method uses Monte-Carlo techniques.
- Thus, the number of iterations N_c does not increase with the number of components.
- Hence, when we have a large number of components, this method is faster than numerical differentiation.
- The computation time of the new algorithm is smaller, but it is still not very fast:
  - the Cauchy method was designed for situations when we can compute the exact value of $p_f(P^{(k)}(A), \ldots)$;
  - in our problem, these values have to be computed by using Monte-Carlo techniques;
  - each such computation itself requires a lot of iterations.

19. How to Decrease the Number of Simulations

- For each combination of values $\delta(A)$, the Monte-Carlo simulation produces an *approximate* value

$$\widetilde{p}_f(\widetilde{P}(A)+\delta_A,\ldots)=p_f(\widetilde{P}(A)+\delta_A,\ldots)+c_n.$$

- The difference c_n is normally distributed with mean 0 and variance $\sigma^2=\dfrac{\widetilde{p}_f\cdot(1-\widetilde{p}_f)}{N}$.
- As a result, the difference $c\stackrel{\text{def}}{=}\widetilde{p}_f(\widetilde{P}(A)+\delta_A,\ldots)-\widetilde{p}_f$ can be represented as $c=c_c+c_n$, where
  - $c_c\stackrel{\text{def}}{=}p_f(\widetilde{P}(A)+\delta_A,\ldots)-\widetilde{p}_f$ is Cauchy distributed,
  - c_n is normal with mean 0 and known σ, and
  - the components c_c and c_n are independent.
- Thus, the characteristic function $\chi(\omega)\stackrel{\text{def}}{=}E[\exp(\mathrm{i}\cdot\omega\cdot c)]$ of c is $\chi(\omega)=\chi_c(\omega)\cdot\chi_n(\omega)=\exp(-|\omega|\cdot\Delta-\omega^2\cdot\sigma^2)$.

20. Resulting Algorithm

- First, we use a lengthy Monte-Carlo simulation to compute the value $\widetilde{p}_f = p_f(\widetilde{P}(A), \ldots)$.
- Then, for $k = 1, 2, \ldots, N$, we repeat the following:
  - use a random number generator to compute n numbers $r_A^{(k)}$ uniformly distributed on $[0, 1]$;
  - compute $\delta_A^{(k)} = \Delta_i \cdot \tan(\pi \cdot (r_A^{(k)} - 0.5))$;
  - use Monte-Carlo simulations to find the frequency (probability estimate) $\widetilde{p}_f(\widetilde{P}(A) + \delta_A^{(k)}, \ldots)$ and then

$$c^{(k)} = \widetilde{p}_f(\widetilde{P}(A) + \delta_A^{(k)}, \ldots) - \widetilde{p}_f.$$

- For $\omega > 0$, compute $\chi(\omega) = \dfrac{1}{N} \cdot \sum_{k=1}^{N} \cos\left(\omega \cdot c^{(k)}\right)$.
- Then, compute $\Delta = -\dfrac{\ln(\chi(\omega))}{\omega} - \sigma^2 \cdot \dfrac{\omega}{2}$.

21. Conclusions and Future Work

- *Problem:*
  - we know lower and upper bounds $\underline{P}(A)$ and $\overline{P}(A)$ on probabilities of component failures $P(A)$;
  - estimate the probability of failure p_f of a complex system such as an aircraft.
- *Assumptions:*
  - failures of different components are independent events, and
  - intervals $[\underline{P}(A), \overline{P}(A)]$ are narrow.
- *Result:* an efficient method that uses Cauchy deviates to estimate the desired range $[\underline{p}_f, \overline{p}_f]$.
- *Future work:* account for possible dependence between component failures.

22. Acknowledgments

- C. Jacob was supported by a grant from @MOST Prototype, a joint project of:
  - Airbus,
  - LAAS-CNRS,
  - ONERA, and
  - ISAE.
- V. Kreinovich was supported:
  - by the Nat'l Science Foundation grants HRD-0734825 and DUE-0926721 and
  - by Grant 1 T36 GM078000-01 from the Nat'l Institutes of Health.