Fault Detection with Uncertain Priors

Yakov Ben-Haim*
Technion—Israel Institute of Technology, Haifa 32000 Israel

In this paper we discuss the design of a failure detection algorithm based on highly uncertain probability distributions for the 'fail' (F) and 'no-fail' (NF) cases. The structure of the algorithm is that NF is declared if and only if the measurement vector x falls in a pre-specified domain D. The designer's task is to choose the domain D so that the probability of missed detection is less than $P_{0\mathrm{c}}$ and the probability of false alarm is less than $P_{1\mathrm{c}}$. Severe uncertainty in the F and NF prior probabilities is represented with info-gap models of uncertainty. The design procedure is developed and the trade-offs between performance (in terms of $P_{0\mathrm{c}}$ and $P_{1\mathrm{c}}$) and robustness to uncertainty are explored. An heuristic example is presented.

I. Theoretical Background

Measurements are made in order to decide whether the measured system is in the no-fail (NF) or fail (F) state. The measurement vector x is a random variable conditioned on the state of the system: NF or F. No-fail is declared if and only if x falls in a pre-specified domain D. The best-estimated (but highly uncertain) probability density functions (pdfs) of x under NF and F are $\widetilde{p}_0(x)$ and $\widetilde{p}_1(x)$ respectively. These estimated pdfs are highly uncertain because of limited data, surprises in operating conditions, variability of the system, unanticipated effects of failure, etc. We use info-gap models to represent this uncertainty[1,2]. An info-gap model is an unbounded family of nested sets of uncertain events, in our case, pdfs. The info-gap models for uncertainty in NF and F are denoted $\mathcal{U}_0(\alpha, \widetilde{p}_0)$ and $\mathcal{U}_1(\alpha, \widetilde{p}_1)$ respectively.

Many info-gap models are available for representing uncertain pdfs. The choice of an uncertainty model depends on the available information about the pdfs. A common info-gap model for uncertain pdfs is the fractional-error model. Let $\mathcal{P}_i$ denote the set of non-negative and normalized pdfs on the domain of $\widetilde{p}_i(x)$. The fractional-error info-gap model for uncertainty in $\widetilde{p}_i(x)$ is:

$$\mathcal{U}_i(\alpha, \widetilde{p}_i) = \left\{ p(x) : \ p(x) \in \mathcal{P}_i, \ |p(x) - \widetilde{p}_i(x)| \le \alpha \widetilde{p}_i(x) \right\}, \quad \alpha \ge 0 \tag{1}$$

The set $\mathcal{U}_i(\alpha, \widetilde{p}_i)$ contains all pdfs allowed at horizon of uncertainty α. These uncertainty-sets become more inclusive as α increases. The horizon of uncertainty, α, is unknown and the info-gap model is an unbounded family of nested sets of possible pdfs. Since the horizon of uncertainty is unknown and unbounded, there is no worst case and we cannot perform a min-max analysis.

Given the pdf $p(x)$ for F, the probability of missed detection is:

$$P_0(p) = \Pr(x \in D | \mathrm{F}) \tag{2}$$

Likewise, given the pdf $p(x)$ for NF, the probability of false alarm is:

$$P_1(p) = \Pr(x \notin D | \mathrm{NF}) \tag{3}$$

We would like to choose D so that $P_0(p) \le P_{0\mathrm{c}}$ and $P_1(p) \le P_{1\mathrm{c}}$. However, we only know the highly uncertain estimated pdfs, $\widetilde{p}_0(x)$ and $\widetilde{p}_1(x)$.

*Yitzhak Moda'i Chair in Technology and Economics, Faculty of Mechanical Engineering, Technion—Israel Institute of Technology, Haifa 32000 Israel.

Given performance requirements P_{0c} and P_{1c}, the robustness of decision-domain D to uncertainty in the estimated pdfs is the greatest horizon of uncertainty α up to which all realizations of the pdfs lead to adequate performance:

$$\widehat{\alpha}(D, P_{0c}, P_{1c}) = \max \left\{ \alpha : \left(\max_{p \in \mathcal{U}_1(\alpha, \widetilde{p}_1)} P_0(p) \right) \le P_{0c}, \quad \left(\max_{p \in \mathcal{U}_0(\alpha, \widetilde{p}_0)} P_1(p) \right) \le P_{1c} \right\} \tag{4}$$

The robustness function, $\widehat{\alpha}(D, P_{0c}, P_{1c})$, is the basis for evaluating the feasibility of performance requirements P_{0c} and P_{1c} and for choosing the decision-domain D.

The main trade-off property is that robustness decreases as demanded performance increases. That is, for missed-detection performance, a smaller (better) value of P_{0c} entails a smaller (worse) value of the robustness to uncertainty in the pdf:

$$P_{0c} < P'_{0c} \quad \text{implies} \quad \widehat{\alpha}(D, P_{0c}, P_{1c}) \le \widehat{\alpha}(D, P'_{0c}, P_{1c}) \tag{5}$$

A similar relation holds for false-alarm performance:

$$P_{1c} < P'_{1c} \quad \text{implies} \quad \widehat{\alpha}(D, P_{0c}, P_{1c}) \le \widehat{\alpha}(D, P_{0c}, P'_{1c}) \tag{6}$$

Furthermore, the robustness vanishes if either of the critical probabilities, P_{0c} or P_{1c}, equals the corresponding estimated probabilites, $P_0(\widetilde{p}_1)$ or $P_1(\widetilde{p}_0)$ respectively. (Recall from eqs.(2) and (3) that $P_0(\widetilde{p}_1)$ is the best-estimate of the probability of missed detection, while $P_1(\widetilde{p}_0)$ is the best-estimate of the probability of false alarm.) Specifically:

$$\widehat{\alpha}(D, P_{0c}, P_{1c}) = 0 \quad \text{if} \quad P_{0c} = P_0(\widetilde{p}_1) \quad \text{or if} \quad P_{1c} = P_1(\widetilde{p}_0) \tag{7}$$

These quantitative trade-offs, eqs.(5)–(7), enable the designer to evaluate alternative domains D in terms of their reliability and performance. Specifically, the robustness function $\widehat{\alpha}(D, P_{0c}, P_{1c})$ induces a preference-ranking on choices of the decision-domain D. More robustness is better than less robustness, at the same level of performance. That is, domain D is preferred over domain D' if the former is more robust at the same values of P_{0c} and P_{1c}:

$$D \succ D' \quad \text{if} \quad \widehat{\alpha}(D, P_{0c}, P_{1c}) > \widehat{\alpha}(D', P_{0c}, P_{1c}) \tag{8}$$

The robust-satisficing domain, at demanded performance P_{0c} and P_{1c}, maximizes the robustness:

$$\widehat{D}(P_{0c}, P_{1c}) = \arg\max_{D} \widehat{\alpha}(D, P_{0c}, P_{1c}) \tag{9}$$

The robust-satisficing domain may differ from a domain which is optimal with respect to the best-estimated pdfs. Let $D^\star$ be a domain which is Pareto optimal in the sense that any change in $D^\star$ which reduces (improves) one of the failure probabilities also increases (degenerates) the other. Denote the best-estimated failure probabilities with $D^\star$ as:

$$\pi_0 = P_0(\widetilde{p}_1, D^\star), \quad \pi_1 = P_1(\widetilde{p}_0, D^\star) \tag{10}$$

From eq.(7) we know that $\widehat{\alpha}(D^\star, \pi_0, \pi_1) = 0$, so that performance as good as expected from $D^\star$ cannot be relied upon: $D^\star$ has no immunity to error in the pdfs. A robust-satisficing choice of D, eq.(9), will be one which achieves adequate (but Pareto-sub-optimal) failure probabilities with maximal immunity to error in the estimated pdfs.

II. Example: Nominally Exponential Distributions

In this section we develop the robustness function for the special case that the best-estimated pdfs for the no-fail (NF) and fail (F) conditions are exponential:

$$\widetilde{p}_i(x) = \lambda_i \mathrm{e}^{-\lambda_i x}, \quad i = 0,\ 1 \tag{11}$$

We will suppose that $\lambda_0 > \lambda_1$, meaning that the nominal NF distribution is concentrated more to the left than the nominal F distribution. The info-gap models for uncertainty in the actual pdfs are the fractional-error models of eq.(1).

The robustness of eq.(4) is the greatest horizon of uncertainty, α, up to which both failure probabilities are less than their critical values. Let $\widehat{\alpha}_i(D, P_{ic})$ denote the robustness with respect to criterion $i = 0$ or $i = 1$, where $j = 1 - i$:

$$\widehat{\alpha}_i(D, P_{ic}) = \max \left\{ \alpha : \left(\max_{p \in \mathcal{U}_j(\alpha, \widetilde{p}_j)} P_i(p) \right) \le P_{ic} \right\} \tag{12}$$

The overall robustness is the lesser of these two values:

$$\widehat{\alpha}(D, P_{0c}, P_{1c}) = \min \left\{ \widehat{\alpha}_0(D, P_{0c}),\ \widehat{\alpha}_1(D, P_{1c}) \right\} \tag{13}$$

We will derive expressions for $\widehat{\alpha}_0(D, P_{0c})$ and $\widehat{\alpha}_1(D, P_{1c})$ and study the properties of $\widehat{\alpha}(D, P_{0c}, P_{1c})$.

Consider a threshold test, so the domain of x values which is declared NF is $D = [0, x_s]$. The designer must choose x_s. Thus eqs.(2) and (3), for the probability of missed detection and false alarm, are:

$$P_0(p) = \int_0^{x_s} p(x)\,\mathrm{d}x, \quad P_1(p) = \int_{x_s}^{\infty} p(x)\,\mathrm{d}x \tag{14}$$

Let $\widetilde{x}_{im}$ denote the median of $\widetilde{p}_i(x)$. We will assume that x_s exceeds the median values of both nominal distributions: $x_s > \widetilde{x}_{0m}$ and $x_s > \widetilde{x}_{1m}$.

At horizon of uncertainty $\alpha \le 1$ one can readily show that the inner maximum in the definition of $\widehat{\alpha}_0(D, P_{0c})$ occurs with the following pdf from $\mathcal{U}_1(\alpha, \widetilde{p}_1)$:

$$p(x) = \begin{cases} (1+\alpha)\widetilde{p}_1(x), & \text{if } x \le \widetilde{x}_{1m} \\ (1-\alpha)\widetilde{p}_1(x), & \text{if } x > \widetilde{x}_{1m} \end{cases} \tag{15}$$

This is a normalized pdf belonging to $\mathcal{U}_1(\alpha, \widetilde{p}_1)$ which maximizes the probability of missed detection at horizon of uncertainty α. Using this pdf and equating $P_0(p)$ to P_{0c} results in:

$$\widehat{\alpha}_0(D, P_{0c}) = \begin{cases} \dfrac{P_{0c} - P_0(\widetilde{p}_1)}{1 - P_0(\widetilde{p}_1)} & \text{if } P_0(\widetilde{p}_1) \le P_{0c} \\ 0 & \text{else} \end{cases} \tag{16}$$

Likewise, for $\alpha \le 1$, one can readily show that the inner maximum in the definition of $\widehat{\alpha}_1(D, P_{1c})$ occurs with the following pdf from $\mathcal{U}_1(\alpha, \widetilde{p}_1)$:

$$p(x) = \begin{cases} (1+\alpha)\widetilde{p}_0(x), & \text{if } x \ge \widetilde{x}_{0m} \\ (1-\alpha)\widetilde{p}_0(x), & \text{if } x < \widetilde{x}_{0m} \end{cases} \tag{17}$$

This is a normalized pdf belonging to $\mathcal{U}_0(\alpha, \widetilde{p}_0)$ which maximizes the probability of missed detection at horizon of uncertainty α. Using this pdf and equating $P_1(p)$ to P_{1c} results in:

$$\widehat{\alpha}_1(D, P_{1c}) = \begin{cases} \dfrac{P_{1c} - P_1(\widetilde{p}_0)}{P_1(\widetilde{p}_0)} & \text{if } P_1(\widetilde{p}_0) \le P_{1c} \le 2P_1(\widetilde{p}_0) \\ 0 & \text{else} \end{cases} \tag{18}$$

Values of $\widehat{\alpha}_1(D, P_{1c})$ for $P_{1c} > 2P_1(\widetilde{p}_0)$ require revision of eq.(17). We will not pursue this here.

The robustness curves of eqs.(16) and (18) are shown schematically by the two straight lines in fig. 1, illustrating the trade-off relations of eqs.(5)–(7): better performance (lower critical failure probability P_{ic}) entails lower robustness (lower $\widehat{\alpha}_i(D, P_{ic})$). Furthermore, the best-estimated failure probabilities, $P_0(\widetilde{p}_1)$ and $P_1(\widetilde{p}_0)$, each correspond to zero robustness. The thick lines portray the overall robustness defined in eq.(13).

Fig. 1 also illustrates the choice of the critical probabilities, P_{0c} and P_{1c}. In this figure:

$$P_0(\widetilde{p}_1) < P_1(\widetilde{p}_0) < 1 - P_0(\widetilde{p}_1) \tag{19}$$

This implies that the robustness curve for missed detection, $\widehat{\alpha}_0(D, P_{0c})$, is less steep than, and intersects the horizontal axis to the left of the robustness curve for false alarm, $\widehat{\alpha}_1(D, P_{1c})$. Recall, from eq.(13), that the

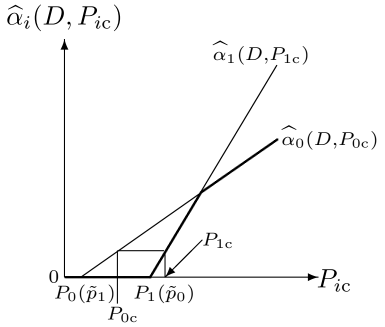

Figure 1: Schematic robustness curves, eqs.(16) and (18).

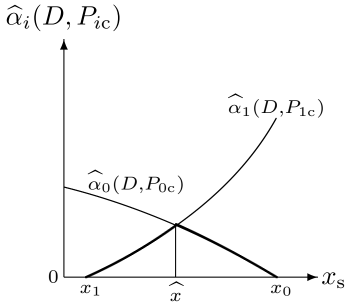

Figure 2: Robustness vs. decision threshold, showing the robust-satisficing threshold $\widehat{x}$.

overall robustness, $\widehat{\alpha}(D, P_{0c}, P_{1c})$, is the lesser of the two robustnesses shown in fig. 1. For any choice of the critical probability of false alarm, P_{1c}, the thin lines show the smallest value of the critical probability of missed detection, P_{0c}, which does not reduce the overall robustness.

The positive parts of eqs.(16) and (18) can be written explicitly as:

$$\widehat{\alpha}_0(D, P_{0c}) = 1 - (1 - P_{0c})e^{\lambda_1 x_s} \quad (20)$$

$$\widehat{\alpha}_1(D, P_{1c}) = P_{1c}e^{\lambda_0 x_s} - 1 \quad (21)$$

For fixed critical probabilities P_{0c} and P_{1c}, the robust-satisficing choice of x_s, which maximizes the robustness as defined in eq.(9), is the value of x_s at which these two expressions are equal. This is illustrated schematically in fig. 2 where:

$$x_0 = -\frac{\ln(1 - P_{0c})}{\lambda_1} \quad (22)$$

$$x_1 = -\frac{\ln P_{1c}}{\lambda_0} \quad (23)$$

The robustness curves cross if $x_0 > x_1$ which occurs in the common case that $1 - P_{0c} > P_{1c}$ and $\lambda_1 < \lambda_0$. The thick lines constitute the overall robustness according to eq.(13).

This example illustrates the management of severe uncertainty in prior probability distributions, when designing a fault detection algorithm. When best-estimated failure probabilities are highly unreliable, the info-gap methodology enables the designer to identify design parameters whose level of performance can be reliably anticipated.

References

1. Yakov Ben-Haim, 2001, 2006, *Information-gap Decision Theory: Decisions Under Severe Uncertainty,* Academic Press, San Diego.

2. Yakov Ben-Haim, 2005, Info-gap Decision Theory For Engineering Design. Or: Why 'Good' is Preferable to 'Best', appearing as chapter 11 in *Engineering Design Reliability Handbook,* Edited by Efstratios Nikolaidis, Dan M.Ghiocel and Surendra Singhal, CRC Press, Boca Raton.