Défaillance de systèmes

Nicolas ROLIN
nicolas.rolin@laposte.net
Maître de stage : Nicolas Vayatis

24 juin 2010

Lieu : CMLA

Table des matières

Introduction

Ce stage s'est fait dans le cadre d'un projet (le projet craft) en collaboration avec le lurpa et le lsv, pour étudier la probabilité de panne d'un système complexe. J'ai donc lu les travaux déjà réalisés dans le projet (dans d'autres domaines que les mathématiques) et essayé de répondre à ce problème avec un autre point de vue, qui est statistique. En effet des résultats étaient déjà présents (calculs analytiques) mais présentaient certains inconvénients (lourdeur des calculs, d'où choix de lois restreints pour mener les calculs). Dans ce cadre j'ai tenté de mettre au point un algorithme permettant de retrouver cette probabilité à l'aide de simulations numériques.

1 Problèmes

Le domaine de la sûreté de fonctionnement concerne l'étude des outils et des méthodes qui permettent de garantir la fiabilité d'un système complexe, sa maintenabilité, sa disponibilité et sa sécurité. Les arbres de défaillance sont utilisés particulièrement pour l'étude de fiabilité des systèmes complexes. Ils permettent de modéliser les relations causales de défaillance d'un système entre les différentes entités basiques qui peuvent provoquer une défaillance globale du système et ainsi estimer la probabilité d'apparition d'une défaillance.

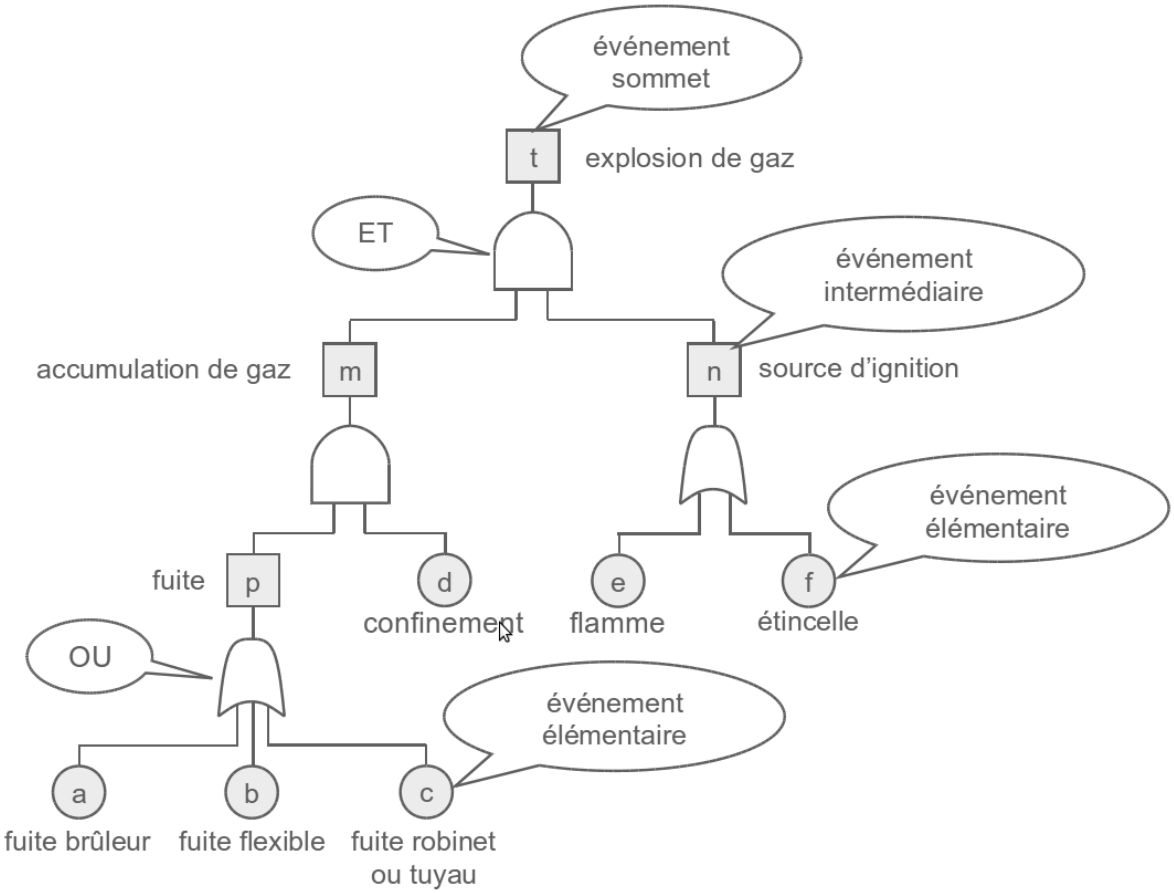

FIG. 1 – Arbre de défaillance d'une bombonne de gaz

1.1 Arbres de défaillance

1.1.1 Arbres de défaillance statiques

Cet exemple (FIG. 1) modélise le fonctionnement d'une bonbonne de gaz. La défaillance globale étudiée, ou *événement sommet*, correspond à l'explosion de la bonbonne de gaz. L'arbre est constitué d'un ensemble d'événements élémentaires (tel que "fuite brûleur", "source d'ignition") ainsi que des relations causales entre ces événements élémentaires.

Par exemple, l'événement sommet est déclenché lorsque deux évènements apparaissent, l'accumulation de gaz et une source d'ignition. La relation logique décrivant cette relation correspond au "ET" logique et est modélisée par une porte "ET" logique. La source d'ignition est elle même provoquée par une flamme ou une étincelle, relation modélisée par une porte "OU" logique. De manière analogue, l'accumulation de gaz résulte d'un confinement et d'une fuite (porte "ET"), qui elle même peut être une fuite brûleur, une fuite flexible ou une fuite robinet ou tuyau (porte "OU").

On a ainsi décomposé un évènement complexe (explosion) en de multiples évènements élémentaires (étincelle,fuite), leur relation étant modélisée par une formule logique (ici composée uniquement de opérateurs "ET" et "OU"). On appellera par la suite la formule logique qui relie les évènements élémentaires à l'évènement sommet la *fonction de structure*.

1.1.2 Arbres de défaillance dynamiques

Les systèmes réels nécessitent parfois de prendre en compte l'ordre dans lequel les évènements surviennent. Par exemple, si une étincelle apparaît puis du gaz confiné, la bonbonne n'explose pas. Par ailleurs, des évènements élémentaires peuvent avoir des lois non indépendantes : un moteur de rechange ne s'use que lorsque le moteur principal est hors service. Ce type de dépendances ne peuvent être exprimé par des portes logiques statiques, leur modélisation nécessite l'utilisation de *portes dynamiques* intégrant une composante temporelle. Nous noterons t_A^i le temps d'apparition de l'évènement A dans la simulation i.

Deux portes dynamiques sont étudiés dans la suite :

- porte **PAND** : cette porte comporte deux entrées A et B, et dont la sortie est vraie si A et B sont vrais et l'évènement A s'est produit avant l'évènement B, i.e $t_A^i < t_B^i$.
- porte **spare** : une porte spare est une porte comportant deux entrées A et B, où A est l'évènement primaire et B l'évènement spare. On la note $A \lhd B$.
  L'évènement B a deux états : dormant (B_d) et actif (B_a). B est de loi B_d jusqu'à t_A, puis de loi B_a . La sortie de la porte est vraie si A et B sont vrais

  Cold spare
  Une cold spare est une porte spare où la loi de B est 0 tant que A n'est pas survenu.

  Warm spare
  Une warm spare (cas général) est une porte spare où la loi de B est différente selon que A survienne où non.

1.1.3 Séquences de coupes minimales

On note $[a_1, a_2, ..., a_n]$ une *séquence* si $a_1, a_2, ..., a_n$ sont des éléments qui sont survenus dans l'ordre $t_{a_1}^i < t_{a_2}^i < ... < t_{a_n}^i$.
Dans le cas des arbres dynamiques, la fonction de structure est donc une fonction qui à une séquence $[a_1, a_2, ..., a_n]$ associe un résultat binaire.
On peut donc s'intéresser à chercher, pour un arbre donné, quelles sont les séquences qui provoquent la panne du système (séquences de coupe).
Si on définit la relation d'ordre sur les séquences $\leq$ par

$$[a_1, a_2, ..., a_n] \leq [b_1, b_2, ..., b_p] \Leftrightarrow \forall 1 \leq i < j \leq n, \exists b_k, \exists b_l k < l, a_i = b_k \text{ et } a_j = b_l$$

On peut donc chercher l'ensemble minimal des séquences de coupes pour $\leq$.
Cela nous permet d'avoir une représentation équivalente et simple de la fonction de structure.

1.2 Objectifs

On souhaite calculer la probabilité que l'évènement sommet a de survenir avant un temps T, quelque soit l'arbre dynamique que l'on étudie.

1.2.1 Calcul analytique sur la fonction de structure

La première méthode consiste à calculer directement la probabilité de l'évènement sommet à partir de la fonction de structure.
Pour cela il faut construire un modèle temporel des opérateurs booléens [3]. On note Pr{A}(t) la probabilité que l'évènement A a de survenir avant le temps t. Pour un évènement élémentaire A, on note f_A la fonction de densité de probabilité de A et F_A la fonction de répartition de A.

On a alors si A et B sont indépendants :

$$Pr\{A \cdot B\}(t) = PrA(t)PrB(t)$$

$$Pr\{A + B\}(t) = PrA(t) + PrB(t) - PrA(t)PrB(t)$$

$$Pr\{A \lhd B\}(t) = \int_0^t f_B(u)F_A(u)\,\mathrm{d}u$$

Si Q est la sortie d'une porte spare avec A en évènement principal, B l'évènement spare (B_d dormant, B_a actif) et $f_a(t, t_A)$ la fonction de densité de probabilité de B actif (qui dépend de B), on a :

$$Pr\{Q\}(t) = \int_0^t (\int_v^t f_{B_a}(u, v)\,\mathrm{d}u) f_A(v)\,\mathrm{d}v + \int_0^t f_A(u)F_B(u)\,\mathrm{d}u$$

Il suffit alors de choisir des distributions pour les évènements élémentaires (on choisit souvent une distribution exponentielle).

L'inconvénient de la méthode est que la complexité du calcul dépend fortement de la nature des lois (en pratique on ne prend que des distributions exponentielles).

1.2.2 Méthode de Monte-Carlo

Une autre méthode consiste à simuler des réalisations du système afin de mesurer expérimentalement l'espérance de l'évènement sommet.
Elle a déjà été appliquée avec succès pour des arbres de défaillance statiques, mais pas encore pour des arbres de défaillance dynamiques [1].
La forme la plus pratique de la fonction de structure pour simuler est la forme de l'ensemble des séquences minimales, disjonction de portes PAND portant sur les entrées.

Bien que donnant un résultat non exact, cette méthode à plusieurs avantages :

- modulable : possibilité de calculer facilement la probabilité d'occurrence de l'évènement sommet en connaissant les probabilités des évènements élémentaires (et leur caractère *spare*), les séquences de coupes minimales (ou la fonction de structure) et le temps considéré ;

- étude de la stabilité : possibilité de savoir si des perturbations sur les paramètres des lois modifient sensiblement le résultat.

2 Modélisation

2.1 Méthodes statistiques

La simulation nous permet de mesurer deux valeurs : la moyenne m et l'écart type σ. On cherche à connaître l'espérance $\mathbb{E}(X)$ de notre variable X (probabilité de tomber en panne). Pour cela on utilise le théorème central limite nous permettant d'avoir des informations sur la distribution :

Théorème 1 ***(Théorème central limite)*** *Soit $(X_n)_{n\leq 1}$ une suite de v.a. i.i.d telle que $\mathbb{E}(X_1^2) < \infty$. On note $\sigma^2 = V(X_1)$. Alors si $m = \mathbb{E}(X_1)$ et $\overline{X_n} = \sum_{i=1,n} X_i/n$:*

$$\sqrt{n}(\overline{X_n} - m) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$$

On cherche à avoir un intervalle de confiance où se situe la moyenne En notant $\overline{\sigma}$ l'écart type empirique, n le nombre de simulations que l'on fait ,ϕ la loi normale centrée réduite, p la probabilité qu'il se trouve dans l'intervalle et k la largeur de l'intervalle, on a alors

$$n \geq \left(\frac{2\overline{\sigma}\phi^{-1}(1/2 + p/2)}{k} \right)^2$$

On cherche à avoir un intervalle de confiance où se situe la moyenne. Pour cela on va utiliser l'inégalité de Hoeffdig :

Théorème 2 ***(Inégalité de Hoeffdig)*** *Soit une suite $(X_k)_{1\leq k\leq n}$ de v.a réelles indépendantes. On pose $S_n = X_1 + X_2 + ... + X_n$. Alors :*

$$\forall t \geq 0, \mathbb{P}\left(|S_n - \mathbb{E}(S_n)| \geq t\right) \leq 2\exp\left(\frac{-2t^2}{n}\right)$$

On a donc

$$\forall t \geq 0, \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| \geq \frac{t}{n}\right) \leq 2\exp\left(\frac{-2t^2}{n}\right)$$

$$\iff \forall t \geq 0, \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| \geq t\right) \leq 2\exp\left(-2t^2 n\right)$$

Si l'on souhaite avoir l'espérance contenue dans l'intervalle $[m - t, m + t]$ avec une probabilité de $1 - \alpha$, il suffit donc de faire $n \geq \frac{ln(\frac{2}{\alpha})}{2t^2}$ simulations

2.2 Algorithme

L'algorithme consiste en la simulation d'un fonctionnement du système sur une durée déterminée T pour savoir si il tombe en panne ou non.
Pour le faire, on divise la simulation en plusieurs étapes :

Calcul des temps d'occurrence des pannes élémentaires :
cas général : On calcule les temps d'occurrence avec le méthode d'inversion.
Soit A un évènement élémentaire de fonction de répartition $F_A(t)$. Soit $F_A^{-1}(u) = \inf\{t/F(t) \geq u\}$.

Lemme 1 *Pour* $0 < u < 1$, $u \leq F(t)$ *si et seulement si* $F^{-1}(u) \leq t$ *[2]*

D'après le lemme 3, si X est une variable aléatoire qui suit une loi uniforme sur $[0, 1]$, alors $F^{-1}(X)$ suit la loi μ de fonction de répartition F.

Exemple 1 *Si la loi de A est exponentielle, on a* $F_A(t) = 1 - e^{-\lambda_A t}$, *d'où* $F_A^{-1}(u) = -\frac{1}{\lambda} log(1-u)$ *Il suffit donc de tirer un nombre dans* $[0, 1]$, *puis de lui appliquer* F_A^{-1} *pour obtenir un temps qui suit la loi de A*

On peut donc facilement simuler les lois qui nous intéressent s'il est possible d'inverser leur fonction de répartition.

Problème des portes spare : La simulation des portes spare se fait en deux temps : on calcule dans un premier temps les temps d'occurrence des évènements A et B_d (A primaire, B spare). Si $t_B < t_A$, on garde le tirage. Sinon on retire le temps t_B selon la loi de B_a

Exemple 2 *Si la loi de* B_a *est exponentielle, on a donc* $F_{B_a}(t, t_A) = 1 - e^{-\lambda_{B_d}(t - t_A}$, *d'où* $F_{B_a}^{-1}(u) = -\frac{1}{\lambda} log(1-u) + t_A$

Calcul des séquences de coupe On commence par comparer le temps d'apparition de chaque évènement avec le temps T, et on ne garde que les évènements A tels que $t_A \leq T$
Puis on classe les évènements restants dans l'ordre chronologique, ce qui nous donne une séquence.

Calcul de l'évènement sommet si la séquence trouvée appartient à l'ensemble des séquences de coupes, alors le système tombe en panne, sinon il fonctionne.

Pour estimer la probabilité de défaillance il suffit de simuler suffisamment de fonctionnement de la machine pour avoir la précision voulue.

2.3 Attentes

On souhaite que notre simulation estime correctement la probabilité voulue, donc que la moyenne de panne pour nos simulations coïncide avec celle calculée analytiquement.
L'écart type permet de vérifier la stabilité de l'estimation calculée.

3 Expérimentation

J'ai utilisé mon algorithme sur deux systèmes différents : l'HCAS (Hypothetical Cardiac Assistant System) et le HCES (Hypothetical Exemple Computer System)

3.1 HCAS

Ce système,(Fig. 2) représentant une pompe cardiaque, est décomposable en 3 partie : le CPU, la pompe et le moteur.[4] Il est constitué d'une porte PAND, d'une cold spare et d'une

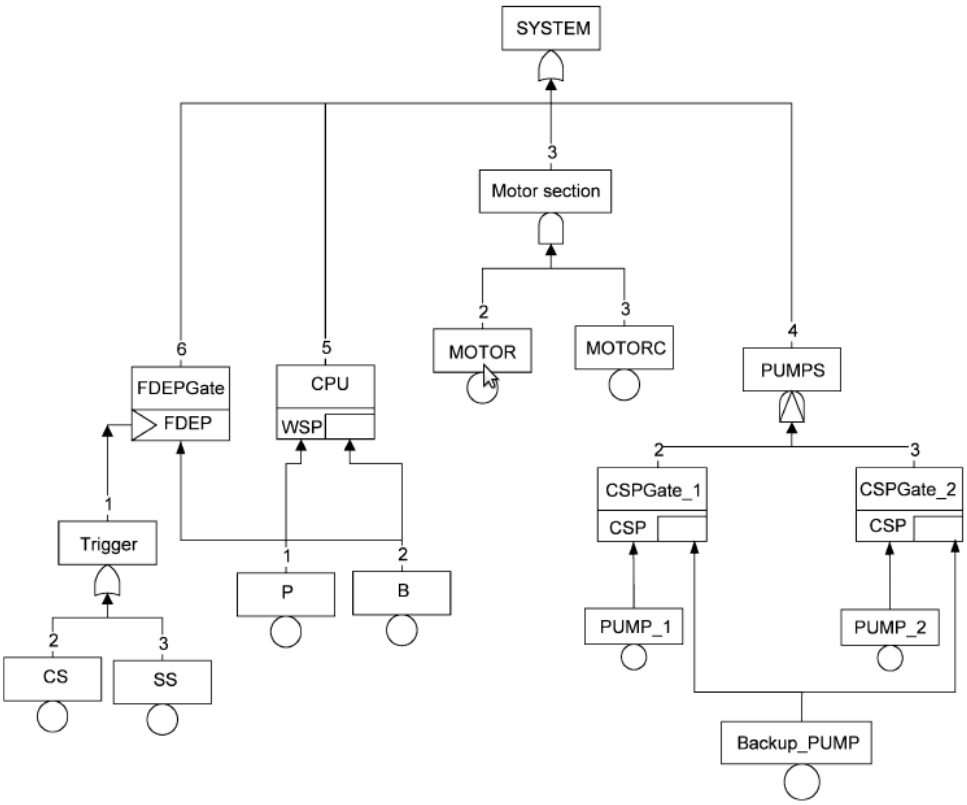

Fig. 2 – Système HCAS

warm spare.
Les résultats présentés (Fig. 2) correspondent à la moyenne de 1000 simulations, constituées elles-mêmes de 500 séquences tirées aléatoirement et indépendamment ; 1000 valeurs sont ainsi calculées, avec l'hypothèse que la distribution est gaussienne autour de la valeur théorique (calculée analytiquement).
Les calculs analytiques ont été menés sur des lois exponentielles (densité $f(t) = le^{-lt}$), et pour des paramètres de l'ordre de 10^{-6} $heure^{-1}$, et donne un résultat de 36,35% de probabilité d'échec La répartition gaussienne semble cohérente. De plus la valeur moyenne est 36,33%, ce qui est bien dans l'intervalle $[\mu - t, \mu + t]$, avec $t = \sqrt{\frac{ln(\frac{2}{\alpha})}{2n}} \approx 0,002302$

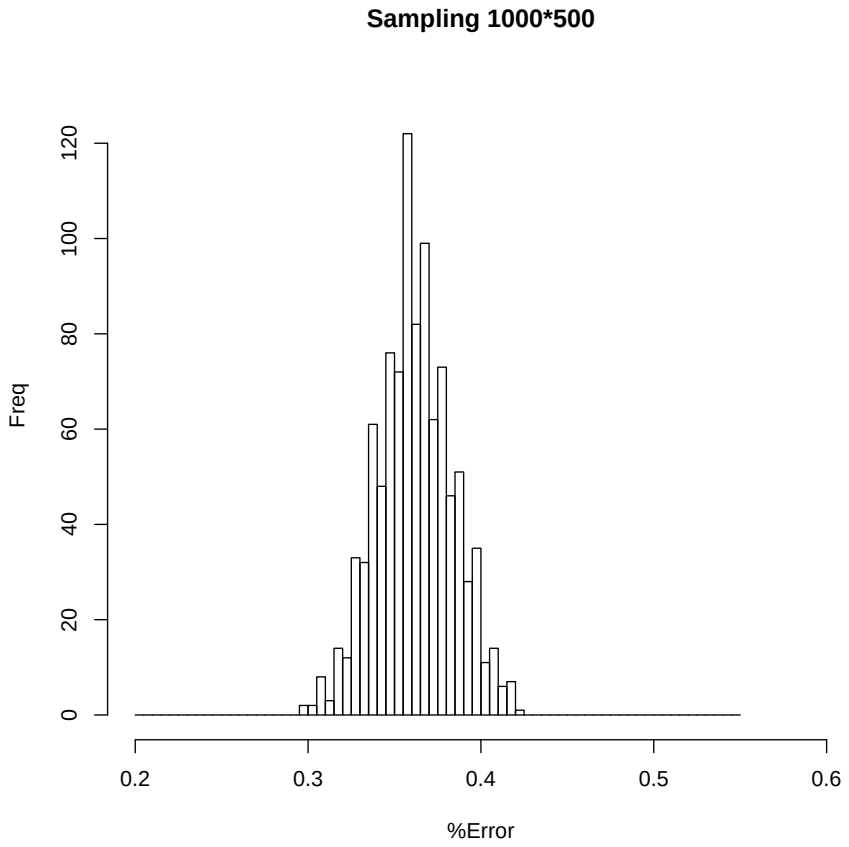

FIG. 3 – répartition des valeurs des simulations

3.2 Discussions

L'hypothèse gaussienne de la distribution est donc vérifiée.

La stabilité est facilement implémentable sur cet algorithme : il suffit de tirer les paramètres dans une distribution choisie avant chaque simulation, puis de continuer l'algorithme normalement. L'augmentation de la complexité n'est pas très importante puisque l'on ne fait que doubler le nombre de tirage aléatoires.

Conclusion

A l'issue de ce stage, j'ai donc mis au point un algorithme qui permet de calculer les probabilités de défaillance de systèmes modélisés par des arbres de défaillance, avec une précision correcte. Il a l'avantage (par rapport au calcul analytique) de ne pas être trop sensible aux lois d'entrée, et de facilement accepter des lois elles-mêmes aléatoires (paramètres non fixes). Il possède néanmoins quelques défauts, comme celui d'être gourmand en nombre de simulation si l'on a des valeurs extrêmes, ce qui pourrait être corrigé en utilisant de l'échantillonnage préférentiel

Références

[1] J. Bank and J. Carson. Discrete-event system simulation. *Prentice-Hall INC*, 1984.

[2] James Antonio Bucklew. *Introduction to Rare Event Simulation.*

[3] G.Merle. *Algebraic modelling of Dynamic Fault Trees, contribution to qualitative and quandtitative analysis.* PhD thesis, 2010.

[4] H.Bouladi and J.B.Dungan. A discrete-time bayesian network reliability modeling and analysis framework. *ELSEVIER*, 2004.