

# Avaliação Probabilística de Risco Aplicação a Sistemas de Geração de Energia Elétrica<sup>1</sup>

G.M. CORDEIRO<sup>2</sup>, C.T. CRISTINO<sup>3</sup>, Departamento de Estatística e Informática, UFRPE, Rua Dom Manoel de Medeiros, s/n Dois Irmãos, 50171-900 Recife, PE, Brasil.

**Resumo.** Neste artigo, serão apresentados os principais conceitos e técnicas utilizadas na Avaliação Probabilística de Risco, com ênfase às aplicações no sistema de geração e transmissão de energia elétrica. O estudo de PRA – sigla de *Probabilistic Risk Assessment*, trabalha com modelagem para sistemas markovianos de estados: dado um conjunto finito  $S$  (equipamentos) para cujos elementos está definido um segundo conjunto também finito de estados, existe uma taxa de transição entre tais estados que indica uma probabilidade de uma certa configuração de  $S$  que é função apenas da conjuração no tempo anterior (no caso discreto). Dois métodos comumente utilizados em análise de PRA são MCS – *Minimum Cut Set*, conjunto de cortes mínimos, e os Diagramas Binários de Decisão (BDD). Outros métodos, tais como redes bayesianas, também são aplicáveis.

**Palavras-chave.** Avaliação Probabilística de Risco; cadeias de Markov; conjunto de cortes mínimos; diagramas binários de decisão; redes bayesianas; risco; taxa de falha.

## 1. Introdução

Avaliação Probabilística de Risco (PRA, sigla da expressão em inglês *Probabilistic Risk Assessment*) é um método de análise compreensiva, estruturada e lógica, que tem por objetivo identificar e avaliar riscos em sistemas tecnológicos complexos, de modo a favorecer ações de segurança e melhoria de desempenho de maneira efetiva [13].

PRA quantifica “métricas de risco”. O termo “métricas de risco” refere-se ao tipo das quantidades que aparecerão em um modelo de decisão: cada objeto com um frequência ou probabilidade de consequências de uma magnitude específica, ou talvez consequências esperadas. Métricas de risco de interesse em geral incluem perda de equipamentos, perda de vidas humanas, ou danos ambientais. Objetivamente, as métricas de risco são definidas a partir de um conjunto de fatores econômicos e políticos. Comumente, e de mais fácil definição, as métricas dizem respeito a perda de grandes equipamentos, cujo valor individual supera a casa de milhões de reais.

---

<sup>1</sup>Pesquisa suportada financeiramente pela CHESF - Cia Hidroelétrica do São Francisco.

<sup>2</sup>gauss@deinfo.ufrpe.br

<sup>3</sup>cristino@deinfo.ufrpe.br

Um dos passos iniciais na análise de risco, fruto do PRA, é a determinação de cenários que são dados em termos da descrição do sistema analisado. O conceito de cenários é formalizado sobre uma configuração topológica de um sistema: a partir da posição relativa de seus componentes, do nível de influência entre tais componentes e de outros aspectos físicos, mais ou menos particulares, determinados pela tecnologia envolvida.

Por definição, o cálculo do risco é feito via uma avaliação quantitativa de probabilidades e impactos. Um procedimento genérico da transformação da avaliação de risco pode ser dada por: (1) coleta de informações; (2) identificação de eventos iniciais; (3) identificação das fontes de risco; (4) avaliação de confiabilidade de sistemas de segurança; (5) avaliação de impacto; (6) avaliação do risco; (7) tomada de decisão.

## 2. Modelos Combinatórios

Os modelos combinatórios expressam o comportamento da falha de sistema de maneira a descrever os modos que as falhas individuais de um componente ou de um subsistema combinam-se para afetar o sistema total. Dois tipos específicos de modelos combinatórios bastante considerados são os modelos de árvore de falha e de digrafos. As árvores de falha têm sido usadas por muito tempo para a análise quantitativa da confiabilidade por causa de sua representação concisa de combinações da falha de sistema. Digrafos, por outro lado, têm sido usados primitivamente para análise quantitativa, identificando fluxos de propagação de erros através de um sistema.

Um digrafo consiste em um conjunto de vértices conectados por arestas dirigidas. Os nós representam a ocorrência de um evento particular, tal como uma falha de um componente individual ou uma ação humana. As arestas descrevem direções causais de propagação de falhas, sugerindo o efeitos de um vértice sobre os outros a ele conectados. Os digrafos podem também incluir as construções da lógica **E** e **OU**. Os modelos do digrafo são razoavelmente fáceis de construir e serrem verificados porque se assemelham geralmente ao diagrama esquemático de um sistema.

Uma árvore de falha é também um modelo baseado em grafos. Tem geralmente um único evento da raiz, no alto da árvore que representa o evento preliminar da falha do interesse. A construção da árvore de falha um modelo é um processo “top-down”, iniciando com uma falha do sistema e seguindo para trás, procurando-se por causas possíveis desta falha. Isto acontece recursivamente até que somente eventos básicos estejam descritos. Um evento básico representa um evento atômico tal como a falha de um único componente. Os eventos básicos são ligados ao evento superior mais próximo por portas de lógica **E** e **OU**. As negações ocorrem somente em árvores de falha não coerentes [7].

Existem trabalhos que se baseiam na geração de conjuntos de corte mínimos (MCS) para um sistema dado. Um MCS é um conjunto dos eventos cuja ocorrência faz com que um evento-alvo específico ocorra, seja ele falha do sistema, seja algum outro evento. Os digrafos e as árvores de falha podem ser ambos usados produzir o conjunto de cortes mínimos para o sistema. Como uma árvore de falha com  $n$

eventos básicos pode ter  $3^n/n$  cortes, este processo é exponencial com respeito a  $n$ .

### 3. Diagramas Binários de Decisão

Lee [11] foi o primeiro a introduzir uma nova representação para os circuitos de chaveamento que viriam ser conhecidos como diagramas binários de decisão (BDD). A simples representação booleana possui limitações. Representações booleanas compõem uma metodologia algébrica e altamente sistêmica, mas também muito inflexível. O BDD é uma tentativa de um caminho para se obter respostas além das limitações da representação booleana. Geralmente, um BDD é um “diagrama” que diz ao usuário como determinar o valor de saída de função lógica examinando-se os valores das entradas. BDDs é uma representação canônica muito compacta de funções booleanas. Os algoritmos de avaliação de tais diagramas que permitem manipular o BDD tem a complexidade proporcional ao tamanho dos grafos que estão sendo operados, e são completamente eficiente, contanto que os grafos não crescer demasiado grande.

O BDD associado a uma fórmula é uma codificação compacta da tabela verdade da fórmula. Esta representação é baseada em decomposição de Shannon [14]. Os teoremas de decomposição (3.1 e 3.2) são o núcleo dos algoritmos de BDD para computar MSC.

**Teorema 3.1** (Decomposição para implicantes primos (Morreale, 1970)). *Seja  $F = v.F_1 + \bar{v}.F_0$  uma fórmula, tal que  $F_1$  e  $F_0$  não dependam de  $v$ . Então o conjunto de primos implicantes de  $F$  é*

$$PI[F] = PI_n \cup PI_1 \cup PI_0,$$

em que  $PI_n = PI[F_1.F_0]$ ,  $PI_1 = \{v.\pi : \pi \in PI[F_1] \setminus PI_n\}$  e  $PI_0 = \{\bar{v}.\pi : \pi \in PI[F_0] \setminus PI_n\}$ .  $\square$

**Teorema 3.2** (Decomposição para MCS). *Seja  $F = v.F_1 + \bar{v}.F_0$  uma fórmula tal que  $F_1$  and  $F_0$  não dependam de  $v$ . seja  $L$  o conjunto de literais relevantes. Então o MCS de  $F$  é descrito com segue:*

- **Caso 1:** *tanto  $v$  quanto sua negação pertencem a  $L$ . Neste caso, o teorema de decomposição é o mesmo que o teorema anterior de decomposição de implicantes primos.*
- **Caso 2:**  *$v$  pertence a  $L$ , mas sua negação não. Neste caso, existem duas maneiras de computar  $MCS[F]$ .*

PRIMEIRA DECOMPOSIÇÃO  $MCS[F] = MCS_1 \cup MCS_0$

onde,  $MCS_0 = MCS[F_0]$  and  $MCS_1 = \{v.\pi : \pi \in MCS[F_1+F_0]/MCS_0\}$ .

SEGUNDA DECOMPOSIÇÃO  $MCS[F] = MCS_1 \cup MCS_0$ ,

onde,  $MCS_0 = MCS[F_0]$ ,  $MCS_1 = \{v.\pi : p \in MCS[F_1] \div MCS_0\}$

$P \div Q = \{\pi \in P \text{ e } \rho \in Q : \rho \not\subseteq \pi\}$

- **Caso 3:** nem  $v$ , nem  $\bar{v}$  pertence a  $L$ . Então  $MCS[F] = MCS[F_1] \cup MCS[F_0]$ .

Sabe-se que qualquer expressão booleana é igual a uma expressão na *forma normal conjuntiva* (FNC) é uma expressão na *forma normal disjuntiva* (FND) a saber:

$$\bigwedge_{j=1}^l \left( \bigvee_{i=1}^{k_j} t_i^j \right) \qquad \bigvee_{j=1}^l \left( \bigwedge_{i=1}^{k_j} t_i^j \right)$$

onde  $\wedge$  representa o operador de conjunção (e),  $\vee$  o operador de disjunção (ou) e  $t_i^j$  são literais.

Em geral, é muito difícil estabelecer quando uma expressão booleana é satisfeita:

**Teorema 3.3** (Cook, 1971). *A satisfação de expressões booleanas é um problema NP-completo.*

Seja  $x \rightarrow y_0, y_1$  o operador *se-então-senão* definido por

$$\begin{aligned} (x \rightarrow y_0, y_1) &= (x \wedge y_0) \vee (\neg x \wedge y_1) \\ &= xy_0 + \bar{x}y_1 \end{aligned}$$

$x$  é chamado *expressão teste*. Todo operador booleano pode ser expresso de tal maneira, por exemplo,  $\bar{x} = \neg x = (x \rightarrow 0, 1)$ ;  $x - -y = x$  implica  $y$  é  $x \rightarrow (y \rightarrow 1, 0)$ ,  $(y \rightarrow 0, 1)$  e, naturalmente,  $x$  é representado por  $x \rightarrow 1, 0$ .

Agora, pode-se definir uma *Forma Normal Se-Então-Senão* (FNS) como a expressão booleana construída inteiramente a partir dos operadores se-então-senão e as constantes 0 (falso) e 1 (verdadeiro), tal que todos testes sejam efetuados somente sobre variáveis [1].

**Teorema 3.4.** *Qualquer expressão booleana é equivalente a uma expressão na forma normal se-então-senão.*

Se para um conjunto de expressões booleanas concatenadas identificam-se as subexpressões de sua FNS correspondente, disse que foi obtido o diagrama binário de decisão (BDD) (Figura 1).

Neste ambiente, um BDD é ordenado se sobre todo caminho do grafo as variáveis respeitam um dada ordem linear  $x_1 < x_2 < \dots < x_n$  e o BDD ordenado é reduzido se (i) nenhum par de nós distintos  $u$  e  $v$  tem a mesma variável sucessora e antecessora, (ii) nenhum nó variável tem sucessor e antecessor distintos.

### 3.1. Modelagem de Atribuição de Risco usando BDD

A conversão de uma árvore de falha em um BDD é baseado nos algoritmos apresentados em [4] e [5]. Além disso para construir o BDD é necessário construir uma

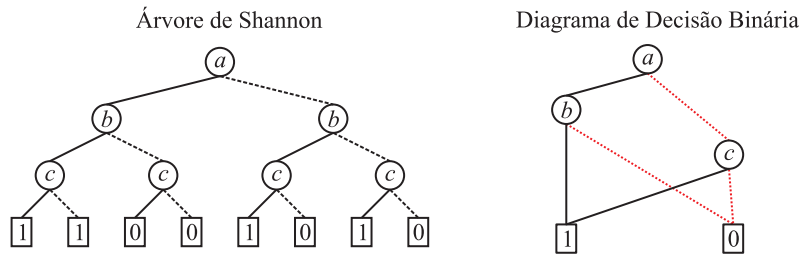


Figura 1: Árvore de decisão e BDD para  $(a \wedge b) \vee (\bar{a} \wedge c)$ .

lista ordenada de nós do sistema que tem um segurança imperfeita. Esta lista pode ser construída como uma árvore de falha, o que estabelece o esquema de ordenação dos eventos.

Cada nó do BDD terá (pelo menos) as seguintes informações associadas a ele:

- O nome do componente.
- Um índice, baseado na ordenação do nó.
- A probabilidade de falha, ou parâmetros que podem ser usados para determinar a probabilidade de falha.
- a probabilidade de segurança para o componente.
- a desconfiança da sub-árvore percorrida para qual o nó esta na raiz. Isto é feito eliminando-se caminhos múltiplos. O valor inicial será 0.
- Ponteiros “esquerda” e “direita” que apontam para o próximo nó no BDD.

## 4. Redes Bayesianas

Uma rede bayesiana é um grafo acíclico dirigido no qual os vértices representam variáveis aleatórias e as arestas influências probabilísticas. As variáveis aleatórias que aparecem nas redes bayesianas representam parâmetros chave que caracterizam o sistema que está sendo modelado. A direção de uma aresta entre dois vértices é geralmente escolhida indicando relações de causa-efeito ou classe-propriedade entre as variáveis denotadas por tais vértices. Em um modelo de rede bayesiana, a distribuição de probabilidade condicional  $P(B|A_1, \dots, A_n)$  é usada para quantificar o peso da influência das variáveis  $A_i$  sobre a variável  $B$ . Os vértices  $A_j$  são chamados os *pais* de  $B$  e  $B$  é chamado um filho de cada  $A_j$ . Uma rede bayesiana representa um modelo probabilístico completo do sistema, pois a distribuição de probabilidade conjunta de qualquer estado elementar do sistema pode ser obtido usando as distribuições de probabilidade condicionais locais e a topologia da rede. Esta propriedade sugere o uso de modelos de redes bayesianas para computar o índice de probabilidade de para avaliação de confiabilidade e risco do sistema elétrico.

A representação de redes bayesianas de um domínio probabilístico também facilita a atualização de probabilidades em resposta a um novo dado, isto é, alguma nova observação sobre o domínio. As probabilidades atualizadas, ou posteriores, são computadas pela propagação de um novo dado por toda rede, observando durante este processo as relações probabilísticas condicionais locais entre as variáveis. Baseados nas probabilidades posteriores, pode-se inferir sobre as causas e/ou efeitos mais prováveis dos fatos observados. Esta característica torna a modelagem por redes bayesianas uma poderosa ferramenta para diagnóstico e análise de predição.

Procedimentos eficientes para a avaliação de propagação probabilística tem sido desenvolvidas [2], [10]. Pacotes de software para construção de redes bayesianas e avaliação computacional estão disponíveis para várias plataformas computacionais e podem ser facilmente utilizadas em uma série de aplicações.

## 5. Atribuição de Probabilidades

Para atribuições de probabilidades para análise de risco, será descrito nesta seção o caso efetuado para o conjunto de geradores da CHESF. Num total de 16 usinas, sendo 14 hidrelétricas e 2 termoeletricas, a CHESF possui 54 geradores. Com o intuito de exemplificar as técnicas de PRA na análise de risco, será colocada a seguinte situação: deseja-se estudar o comportamento futuro dos geradores desta empresa, que relaciona-se ao fornecimento de carga ao sistema e que pode ser um evento de risco à rede.

A partir do histórico de cada gerador, que apresenta todos os eventos ocorridos para cada equipamento, é traçada uma linha de tempo, que indica períodos de funcionamento, falha, manutenção, etc. No caso da CHESF, foram identificadas 43 estados possíveis, tais como manutenções programadas, falhas graves, parada para atendimento ao serviço de transmissão, entre outras. Uma característica óbvia de tal linha é que ela sempre passa do estado de não funcionamento, para o estado de funcionamento. Por exemplo, não se tem manutenção preventiva após uma parada por falta grave. A partir das linhas de tempo, pode-se gerar os grafos que representam os estados (vértices) e suas transições (arestas), com pesos  $\lambda_{ij}$  representando a taxa de transição do estado  $i$  para o estado  $j$  (Figura 2).

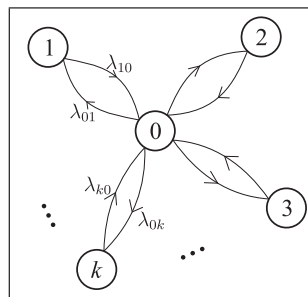


Figura 2: Grafo-estrela que representa os estados de um gerador e suas transições. O vértice rotulado com o estado 0, representa o estado em funcionamento.

Neste ponto, deseja-se estimar os valores dos  $\lambda_{ij}$ 's a fim de se obter um autômato simulador de cenários futuros do comportamento das usinas. Efetivamente os  $\lambda$ -parâmetros são obtidos por

$$\lambda_{ij} = \frac{n_{ij}}{\sum_i t_i},$$

em que  $n_{ij}$  é o número de transições do estado  $i$  para o estado  $j$  e  $\sum_i t_i$  é a soma dos tempos dos intervalos na linha de tempo rotulados com o estado  $i$ . Tais parâmetros, são entrada para o algoritmo que definirá o comportamento futuro do equipamento (Tabela 1).

Tabela 1: Algoritmo para o grafo de estados.

0	Comece no estado $i_0$ ; $n = 0$ ;
1	Estando em um estado $i$ , sorteie para cada aresta $(i, j)$ um tempo $t_{ij} > 0$ aleatório com distribuição exponencial de parâmetro $\lambda_{ij}^{\text{equipamento}}$
2	Permaneça $t_{ij}$ no estado $i$ e vá para o estado $j$ , tal que  $t_{ij} = \min\{t_{ij'} : (i, j') \text{ é aresta.}\};$  $n = n + 1$
3	Se $n < N + 1$ , então volte para 2, senão termine.

Denota-se por  $t_i$  o tempo de permanência do equipamento no estado  $i$ . Se  $T_{ij}$  denotar o tempo de permanência do equipamento no tempo  $i$  até passar para o estado  $j$ , tem-se que

$$T_{ij} \sim E(\lambda_{ij})$$

Neste caso, tem-se que a função de distribuição acumulada para este modelo é

$$\Pr(T_{ij} < x) = 1 - \exp(-\lambda_{ij}x), \quad x > 0$$

cujas funções de densidade de probabilidade são dadas por:

$$f_{T_{ij}}(x) = \lambda_{ij} \exp(-\lambda_{ij}x), \quad x > 0.$$

O tempo de permanência do equipamento do estado  $i$ ,  $T_i$  é:

$$T_i = \min_{j \neq i} \{T_{ij}\}.$$

Da função de distribuição acumulada da distribuição exponencial, tem-se:

$$T_i \sim E(\lambda_i),$$

em que

$$\lambda_i = \sum_{j \neq i} \lambda_{ij}.$$

Para uma estimação no longo horizonte de  $T$  pode, fazendo  $\bar{t}_i$  média dos  $t_i$ , num longo horizonte qualquer  $T$ :

$$\bar{t}_i = \frac{\sum t_i}{n_i}$$

e

$$\hat{\lambda}_i = \frac{1}{\bar{t}_i} = \frac{n_i}{\sum t_i}.$$

Se  $P_i$  é a probabilidade assintótica do equipamento estar no estado  $i$  e  $f_{ij}$ , a frequência de visitas do estado  $i$  para o estado  $j$ :

$$f_{ij} = P_i \lambda_{ij}.$$

Logo, para a estimação no longo horizonte  $T$

$$\hat{P}_i = \frac{\sum t_i}{T} \quad (5.1)$$

$$\hat{f}_{ij} = \frac{n_{ij}}{T}. \quad (5.2)$$

Daí,

$$\hat{\lambda}_{ij} = \frac{n_{ij}}{T} \frac{T}{\sum t_i} = \frac{n_{ij}}{\sum t_i}.$$

A frequência de transições do estado  $i$ ,  $f_i$ , é dada por

$$f_i = P_i \sum_{j \neq i} \lambda_{ij}.$$

E a estimação no longo horizonte  $T$ :

$$\hat{f}_i = \frac{\sum t_i}{T} = \frac{\sum_{j \neq i} n_{ij}}{\sum t_i} = \frac{\sum_{j \neq i}}{T}.$$

O número de transições do estado  $i$  no tempo  $t$  se relaciona com a probabilidade de  $r$  transições do estado  $i$  no tempo  $t$ , para  $r = 1, 2, 3, \dots$ :

$$N_i(t) \sim P(P_i \lambda_i t),$$

em que  $\lambda_i = \sum_{j \neq i} \lambda_{ij}$ .

Agora, se  $N_{ij}(t)$  é o número de transições do estado  $i$  para o estado  $j$  no tempo  $t$ , tem-se que

$$N_{ij}(t) \sim P(P_i \lambda_{ij} t).$$

A probabilidade de  $r$  transições do estado  $i$  para o estado  $j$  no tempo  $t$ , para  $r = 0, 1, 2, \dots$  é dada por

$$\Pr\{N_{ij}(t) = r\} = (r!)^{-1} e^{-P_i \lambda_{ij} t} (P_i \lambda_{ij} t)^r$$

$$\Pr\{N_{ij}(t) > r\} = 1 - \Pr\{\chi_{2(r+1)}^2 > 2P_i \lambda_{ij} t\}.$$



## 6. Conclusões

Neste artigo foram apresentadas algumas técnicas de atribuição probabilística de risco, que correspondem a uma importante ferramenta de tomada de decisão gerencial de diversos sistemas tecnológicos. A partir da clássica definição de risco (dada como probabilidade  $\times$  custo), são efetuados passos na direção de tal conceito. Primeiramente, tem-se a descrição da topologia do sistema, onde são levadas em consideração as influências físicas entre os diversos componentes. Após os eventos de falha serem localizados e classificados, é feita a análise de propagação via funções lógicas, representadas por diagramas de decisão ou por redes bayesianas. As probabilidades dos estados de um equipamento ou de um conjunto deles podem ser obtidas via o autômato definido na Tabela 1.

**Abstract.** In this article, the main concepts and techniques used in the Probabilist Risk Assessment (PRA) will be presented, with emphasis to the applications in the system of generation and transmission of electric energy. The PRA works with modeling for Markov systems of states: given a finite set  $S$  (of equipments) for whose elements we define a also finite set of states, there exist a transition rate between such states that indicates a probability of a certain configuration of  $S$ . This rate is function only of the plot in the previous time (in the discrete case). Two methods comumente used in PRA analysis are Minimum Cut Set (MCS), set of minimum cuts, and the Binary Diagrams of Decision (BDD). Other methods, such as bayesianas networks, also are applicable.

## Referências

- [1] H.R. Andersen, An introduction to binary decision diagrams, *Lecture Notes for 49285 Advanced Algorithms E97*, Department of Information Technology, Technical University of Denmark, 1997.
- [2] S.K. Andersen, K.G. Olesen, F.V. Jensen, F. Jensen, HUGIN - a shell for building Bayesian belief universes for expert systems, em “Proceedings of the Eleventh International Joint Conference on Artificial Intelligent” (AAAI Press, Menlo Park, CA. 1989), pp. 1080.
- [3] R. Billinton, R.N. Alan, “Reliability Evaluation of Power Systems”, Plenum Press, New York, 1984.
- [4] O. Coudert, J.C. Madre, Fault tree analysis:  $10^{20}$  prime implicants and beyond, in “Proceedings of the Reliability and Maintainability Symposium”, pp. 240–245, 1993.
- [5] O. Coudert, J.C. Madre, Towards an interactive fault tree analyser. in “Proceedings of the IASTED International Conference on Reliability, Quality Control and Risk Assessment”, 1992.
- [6] G.M. Cordeiro, C.T. Cristino, E.O.L. Lima, S.B. Melo, Modelos Markovianos, percolação e modelagem em sistemas com grande número de componentes, *Revista de Matemática e Estatística*, **25**, No. 1 (2007), 99–116.

- [7] C.T. Cristino, G.M. Cordeiro, S. Lins, Confiabilidade e avaliação de risco no sistema elétrico, *Revista de Matemática e Estatística*, **25**, No. 2 (2007), 87-109.
- [8] J.L. Doob, “Stochastic Processes”, John Wiley & Sons, New York, 1990.
- [9] W.R. Gilks, S. Richardson, D.J. Spiegelhalter, “Markov Chain Monte Carlo in Practice”, Chapman & Hall, London, 1997.
- [10] S.L. Lauritzen, D.J. Spiegelhalter, Local computations with probabilities on graphical structures and their applications to expert systems, *Journal of the Royal Statistical Society*, **50**, No. 2 (1988), 157–224
- [11] C.Y. Lee, Representation of switching circuits by binary decision programs. *The Bell System Technical Journal*, **38**, (1958), 985–999.
- [12] W. Li, “Risk Assessment of Power Systems - Models, Methods and Applications”, IEEE Press, Piscataway, 2005.
- [13] NASA, “PRA Procedures Guide for NASA Manager and Practitioners”, Version 1.1, em <http://www.hq.nasa.gov/office/codeq/doctree/praguide.pdf>, Washington, 2002.
- [14] A. Rauzy, Mathematical foundations of minimal cutsets. *IEEE Transactions On Reliability*, **50**, No. 4, (2001), 389–396.
- [15] D.C. Yu, T.C. Nguyen, P. Haddawy, Bayesian network model for reliability assessment of power systems, *IEEE Transactions on Power Systems*, **14**, No. 2 (1999), 426–432.