VoI-Based Optimal Sensors Positioning and the Sub-Modularity Issue

Seyed Mojtaba Hoseyni[1], Francesco Di Maio[1], Enrico Zio[1,2,3]
[1]Energy Department, Politecnico di Milano, Milan, Italy
[2]Centre de Recherche sur les Risques et les Crises (CRC), MINES ParisTech/PSL Université Paris, Sophia Antipolis, France
[3] Eminent Scholar at the Department of Nuclear Engineering, Kyung Hee University, Seoul, South Korea
e-mail: seyedmojtaba.hoseyni@polimi.it, francesco.dimaio@polimi.it, enrico.zio@polimi.it

***Abstract*—The problem of sensor positioning for condition monitoring of Systems, Structures and Components (SSCs) has been recently proposed to be addressed by Value of Information (VoI) optimization. VoI is a metric that quantifies the benefit of taking a measurement prior to adopting it. This metric lacks the characteristics of sub-modularity, i.e. the benefit of adding a measurement to a small set of measurement is higher than adding it to a bigger set. This causes the VoI optimization to not guarantee optimal results when the problem is solved by greedy optimization algorithms. In this work, the sub-modularity issue is considered with reference to the thickness gauge sensor positioning on a Steam Generator (SG) of Nuclear Power Plant (NPP), and ways forward to overcome the sub-modularity issue are suggested.**

Keywords-Value of Information; Sensor Positioning; Greedy Optimization; Sub-modularity; Steam Generator.

I. INTRODUCTION

The health condition of Systems, Structures and Components (SSCs) can be monitored by positioning sensors that measure specific properties. Locations where these sensors are to be placed are usually recommended by norms [1]. Recently, sensor positioning has been proposed to be addressed by solving an optimization problem that maximizes the Value of Information (VoI) metric [2-7] by Bayesian decision theory [8,9].

VoI is used to quantify the benefit of taking a measurement regarding specific properties of the SSC in one location rather than another. A measurement is more beneficial than another if it has larger VoI value. Therefore, VoI provides a clear indication on the benefits gained by positioning sensors in specific locations [10].

In principle, given n sensors and j possible sensor locations, to find the optimal subset of locations that maximizes VoI, all possible combination of the n sensors in all candidate locations j are to be evaluated. This optimization problem brings complexity that dramatically increases when n and j increase (e.g., n=30 sensors in j=160 candidate locations, imply 2.74e32 VoI evaluations to be performed) [11]. To overcome this problem, greedy optimization methods have been proposed that consist in looking for the optimal positioning of the sensors by iteratively adding one single sensor, until the optimum arrangement is achieved. By so doing, the number of evaluations reduces to $n \times j$ evaluations (e.g., n=30 sensors in j=160 candidate locations, imply 30×160 VoI evaluations to be performed) [12,13].

Greedy optimization can guarantee optimal or near optimal results, for metrics that satisfy the characteristics of sub-modularity [11], that is, the larger the already identified set the lower the benefit of adding new measurements. For example, the metric M is sub-modular when its increase when it is evaluated at both locations a and b is smaller than when it is evaluated only at location b, (i.e. $[M(a\&b)- M(a)] \leq M(b)$). Incidentally, the VoI metric is not sub-modular, as we shall see in what follows, making the VoI-based greedy optimization for sensor positioning not effective [14].

For engineering discussion purposes, a case study is adopted regarding the positioning of sensors in a Steam Generator (SG) of a Nuclear Power Plant (NPP) that is degrading under creep. The remainder of the paper is as follows: in Section II, a brief description of VoI and optimization problem by the greedy approach are given; in Section III, the sub-modularity is discussed in detail; Section IV presents the case study and the evidence that VoI metric does not have the sub-modularity property; in section V, concluding remarks are given and ways forward to overcome the shortcoming of sub-modularity issue are presented.

II. VOI-BASED GREEDY OPTIMIZATION

Let F be a spatial model of a system where each physical characteristic can be measured by random quantities $f(\bar{x})$ at each location $\bar{x}$ resulting in a vector $\bar{f}(\bar{x})$ representing a multivariate random field in which the various characteristics are jointly described. Vector $\bar{y}$ consists in a subset of measurements $\bar{Y}$ that can be used to update the prior distributions of the multivariate random field $\bar{p}_F$ in the posterior distributions $\bar{p}_{F|\bar{y}}$, by the Bayesian inference method [15,16]. Based on the distributions, the decision maker can take actions $\bar{a}$ to counteract the degradation progression of the SSC that can result in a loss (i.e., a negative utility) described by a loss function $L(\bar{f}, \bar{a})$.

The decision can either be taken based on the prior knowledge $\bar{p}_F$ by minimizing the prior expected loss:

$$\mathbb{E}L(\emptyset) = \min_{\boldsymbol{a} \epsilon A}\{\mathbb{E}_F L(\bar{f}, \bar{a})\} \tag{1}$$

or on the random field properties updated by the Bayesian inference $\bar{p}_{F|\bar{y}}$, whose expected loss is:

$$\mathbb{E}L(\bar{y}) = \mathbb{E}_{\bar{Y}} \min_{\boldsymbol{a} \epsilon A}\{\mathbb{E}_{F|\bar{y}} L(\bar{f}, \bar{a})\} \tag{2}$$

The benefits of making decisions supported by the information acquired, namely based on $\bar{p}_{F|\bar{y}}$ are measured by the VoI, which is calculated as the difference between the prior and posterior expected loss as:

$$VoI\ (\bar{y}) = \mathbb{E}L(\emptyset) - \mathbb{E}L(\bar{y}) \tag{3}$$

VoI quantifies the loss reduction (i.e., the benefit) of decisions taken based on the posterior $\bar{p}_{F|\bar{y}}$, when the set of measurements $\bar{y}$ has been collected. Taking measurements in different locations could bring different effects on the posterior and, ultimately, on the benefits on decision making. Therefore, measurement source locations (i.e., sensors positioning) are to be optimally identified. One strategy might be to simulate monitoring sensors on each location $\bar{x}$ and quantifying the consequent VoI: the maximum VoI would correspond to the optimal positioning for collecting the optimal set of measurements $\bar{Y}^*$. This can be done by solving the optimization problem in (4):

$$\bar{Y}^* = argmax_{\bar{Y} \subseteq \Omega_Y}\ (VoI(\bar{Y}) - C(\bar{Y})) \tag{4}$$

where $C(\bar{Y})$ is the cost of measurement and Ω_Y is the spatial domain where the j locations for taking measurements are assumed to be.

This is, in principle, manageable if the number of sensors n and the number of candidate locations j are limited. In this case, greedy optimization can be used to solve the optimization problem iteratively, assuming that sensor locations are added one at a time: first, the optimal location is found for one sensor; then, another sensor is added considering $\bar{p}_F$ as the previous iteration's posterior $\bar{p}_{F|\bar{y}}$; the process is continued iteratively until the optimal set of measurements $\bar{Y}^*$ is found. In other words, the set of optimal measurements $\bar{Y}^*$ comprises of n sensors that are selected individually, and not as a number of n sensors optimally selected as a whole.

III. SUB-MODULARITY

The greedy optimization correctness requires the solution of the optimization problem to be based on a metric that is sub-modular [17]. In general terms, a sub-modular metric implies that when a specific measurement $\bar{y}$ is added to a measurement set Y, its benefit decreases as the size of the measurement set Y increases. In mathematical terms, let $\bar{Y}_1$ to be a smaller set of measurements than $\bar{Y}_2$ and $\bar{y}$ a specific measurement added to both these sets. Metric M is sub-modular when for every $\bar{Y}_1 \subseteq \bar{Y}_2 \subseteq \Omega_Y$, $\bar{y} \in \Omega_Y$ and $\bar{y} \notin \bar{Y}_2$:

$$M(\bar{Y}_1 \cup \bar{y}) - M(\bar{Y}_1) \geq M(\bar{Y}_2 \cup \bar{y}) - M(\bar{Y}_2) \tag{5}$$

In other words, sub-modularity implies a "diminishing return" property, meaning that when a new set $\bar{y}$ is added to a smaller set $\bar{Y}_1$, the metric improves more than when the same set $\bar{y}$ is added to a bigger set $\bar{Y}_2$ (which also includes the smaller set $\bar{Y}_1$).

VoI is a metric that does not satisfy the sub-modularity property [11]. This means that when a measurement set $\bar{y}$, that may consist of measurements at one or more sensor locations, is added to another measurement set Y, there is no guarantee that the VoI value is less than (or equal to) the VoI value for the set $\bar{y}$ alone. Therefore, when solving the VoI-based optimization problem by greedy optimization, which individually selects single locations to build up the optimal set of locations, there is no guarantee to find the optimal solution. Indeed, since information taken by one sensor is different from that taken from another sensor, it is possibile that a sensor that might be individually selected by the greedy approach because bringing the max VoI, may not be selected when considering a combined set of sensors since the combination of the whole set of sensors not only covers the information of the single sensor but may also cover other sensors in locations with high importance and higher VoI, as it will be shown in the case study.

IV. CASE STUDY

A manifold of the SG of a Prototype Fast Breeder Reactor (PFBR) whose geometric characteristics (Length L and outer diameter D_e) are shown in Fig. 1 is selected as the case study, to show the sub-modularity issue that VoI-based greedy optimization suffers from, when used for optimal sensor positioning. The manifold is here modeled as a spatially distributed rectangular plate Ω_x of length L and width πD_e, whose surface is discretized into 160 squares of 100×100 mm. it is assumed that there is a circumferential welding at the middle of the manifold. Ultrasonic thickness gauges that measure the thickness of base material can be placed at the center of the squares (shown by bold circles in Fig. 1, i.e., j = 160 candidate locations where sensors can be placed).

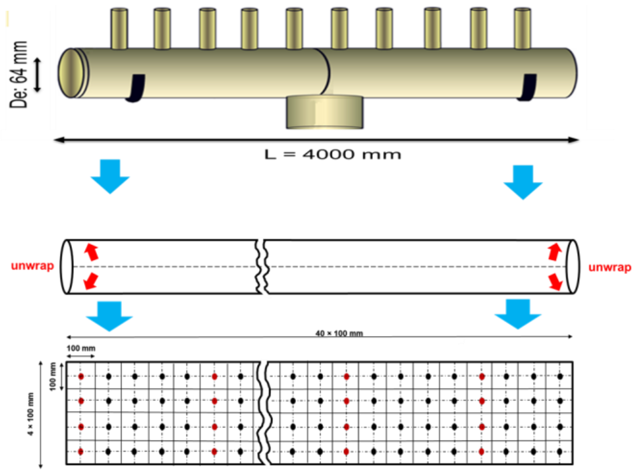

Figure 1. Sketch of the manifold of the SG and its spatially distributed model Ω_x

The thickness model of the manifold $f(\bar{x})$ is assumed to be a Gaussian process model where thickness of each location is normally distributed $\mathbb{N}$~(20mm, 1mm), except for the welding area and adjacent Heat Affected Zones (HAZ), where the uncertainty on thickness is larger and, therefore, the Gaussian process is assumed to be $\mathbb{N}$~(20mm, 2mm). Fig.2 shows the standard deviation of the thickness random field where the majority of the manifold has 1 mm standard deviation except the welding and HAZ area in center that has 2 mm standard deviation.

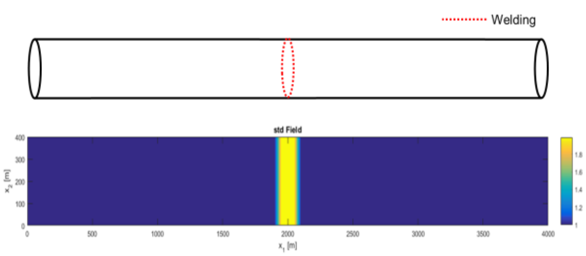

Figure 2. Standard deviation of the thickness random field

The lower the thickness of the SG is, the larger is the probability of failure due to creep. The failure threshold at this SG operational conditions is taken equal to 16.9 mm [18].

Since the decision maker is unaware of the true state $\bar{s}$ of the manifold (i.e., failed (s = 0), operational (s = 1)), based on his/her prior knowledge $\bar{p}_F$, he/ she can make decision about the so-called prior actions $\bar{a}$: 1. do nothing to counteract creep escalation (a = 0) or 2. preventing failure (a = 1) for example by weld repair or reducing operational load. However, if a set of measurements $\bar{y}$ is taken, the random field $\bar{p}_F$ can be updated to the posterior field $\bar{p}_{F|\bar{y}}$, that represents posterior knowledge that supports the decision maker in deciding on so-called posterior actions. The loss function that relates the actions and states with the random field characteristics is defined as:

$$L(\bar{f},\bar{a}) = \begin{cases} 0 & If\ s = 1\ and\ a = 0 \\ C_f = 200K€ & if\ s = 0\ and\ a = 0 \\ C_p = 5K€ & if\ a = 1 \end{cases} \quad (6)$$

where C_f and C_p are the costs of failure and failure prevention, respectively.

A. *Greedy optimization solution for optimal sensor positioning*

Sensors can be placed in j = 160 candidate locations. In this case study, the objective is to position n sensors in every one of 160 locations.

The greedy optimization method includes calculating the prior expected loss $EL(\emptyset)$ of (1) and the posterior expected loss $EL(\bar{y})$ of (2) to quantify the VoI of (3), by assuming placing sensors one at a time in their optimal location. The first location resulting in the max VoI is then identified as shown in Fig. 3 (red circle).

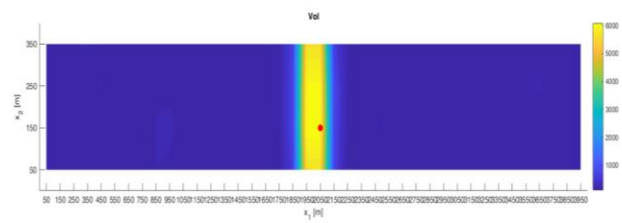

Figure 3. Contour of VoI

The procedure continues iteratively, until the VoI improvements are less than a predefined negligible benefit (here arbitrarily taken equal to 4000 in arbitrary units of costs), resulting in the set of 5 sensors locations, shown in Fig 4. It must be pointed out that even though the number of possible arrangements for n=5 sensors in j=160 locations is 8.2e8, by using the greedy optimization, the VoI evaluation number has been reduced to 800 to find this solution (i.e., 5 iteration× j=160 candidate locations = 800 VoI evaluations).

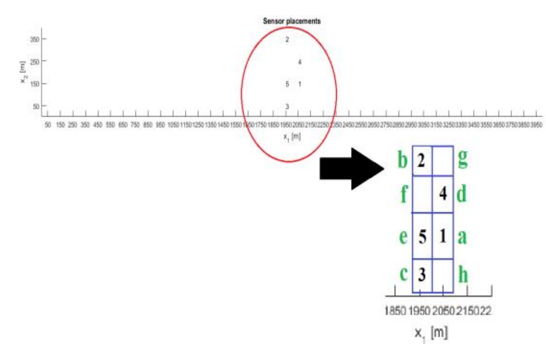

Figure 4. Sensor positioning by greedy optimization

It is worth mentioning that, as expected, all the n=5 sensors are placed in locations with larger uncertainty on the thickness (labeled with *a* to *h* as in Fig. 4): this is to say that acquiring information under this condition provides the highest benefit (i.e., VoI), because it allows reducing the posterior uncertainty. The detailed VoI value of these single sensor locations are presented in Table I.

TABLE I. VOI FOR SINGLE SENSOR LOCATIONS

	Sensors							
	a	***b***	***c***	***d***	***e***	***f***	***g***	***h***
VoI	6203	5691	5693	6132	6177	6091	5701	5682

The sensor location labeled as *a* has the highest VoI value as a single sensor configuration and is selected as the first sensor location by greedy optimization. The values of the VoI in the step by step iterations followed by the greedy method are given in Table II. As it can be seen, since the VoI improvements in the sixth iteration are less than 4000 (i.e., 30043-26495=3548), the first *n*=5 sensors locations are selected as the final optimal set.

TABLE II. VOI FOR GREEDY SENSOR POSITIONING IN EACH ITERATION

	Sensors					
Iteration	***1***	***2***	***3***	***4***	***5***	***6***
Location	***a***	***a, b***	***a, b, c***	***a, b, c, d***	***a, b, c, d, e***	***a, b, c, d, e, f***
VoI	6203	12463	17745	22464	26495	30043

B. The sub-modularity issue

In this Section, we practically show the sub-modularity issue in the case study. As shown in Table II, when at the second iteration, location *b* is added to *a*, the resulting benefit (i.e. {a,b}-{a}=12463- 6203 = 6260 in VoI units) is higher than the VoI value of *b* alone (shown in Table I (i.e. 5691)). This implies that the inequality of (5), with $\bar{Y}_1=\emptyset$ $\bar{Y}_2=\{a\}$ and $\bar{y}=\{b\}$, is not satisfied since adding *b* to *a* results in an increasing return, rather than in a diminishing return (see Fig. 5).

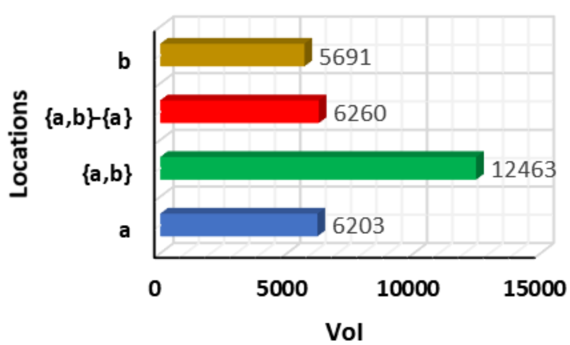

Figure 5. Sub-modularity issue with a sbset od locations

This shows that the VoI metric is not sub-modular, causing the sensor positioning not to be optimum (but still, sub-optimal). If we rely on a non-greedy optimization method such as Particle Swam Optimization (PSO) to optimize the objective function of (4) constrained by the requirement that in *j*=160 candidate locations only a set of *n*=5 sensor locations are to be selected, a different result is obtained as shown in Table III.

TABLE III. VOI OF 5 SENSOR LOCATIONS BY GREEDY AND NON-GREEDY METHODS

Optimization type	Greedy	Non-Greedy
Sensors selected	*a, b, c, d, e*	*a, b, c, f, g*
VoI	26495	26926

The non-greedy optimization results in a VoI 2% larger (26926) than that of the greedy optimization result (26495), since locations *f* and *g* are selected instead of *d* and *e*. Fig. 6 compares the VoI values obtained with the non-greedy optimization (dot) and those obtained by greedy optimization adding one location at a time (line).

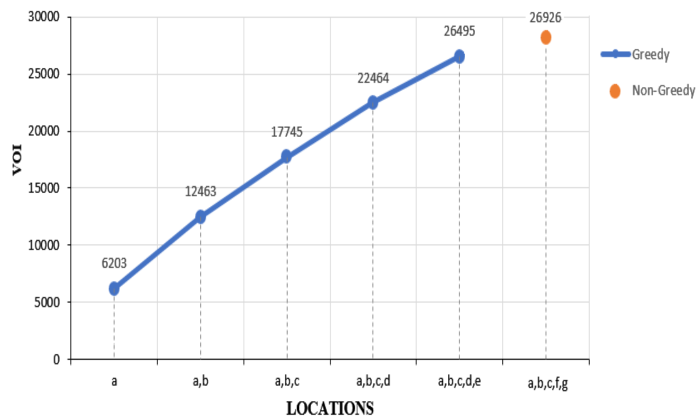

Figure 6. Comparison of the VoI values in greedy and non-greedy optimizations

V. CONCLUDING REMARKS

In this work, the sub-modularity issue of the VoI metric and challenges in its optimization by greedy approaches have been discussed. A case study of a FPBR-SG manifold is presented whose aim is to find the optimal locations of the sensors that measure its thickness reduction, to avoid failure due to creep. The lack of the sub-modularity property of the VoI metric is practically shown and a non-greedy optimization method (PSO) is used to show that this approach can yield better results than greedy approach.

ACKNOWLEDGMENT

The work is developed within the research project "SMART MAINTENANCE OF INDUSTRIAL PLANTS AND CIVIL STRUCTURES BY 4.0 MONITORING TECHNOLOGIES AND PROGNOSTIC APPROACHES - MAC4PRO ", sponsored by the call BRIC-2018 of the National Institute for Insurance against Accidents at Work – INAIL (CUP: J56C18002030001).

REFERENCES

[1] M. H. Faber and R. Rackwitz, "Sustainable decision making in civil engineering," *Structural Engineering International* , vol. 14, no. 3, pp. 237-242, 2004.

[2] C. Malings and M. Pozzi, "Value of Information for Spatially Distributed Systems: application to sensor placement," *Reliability Engineering & System Safety,* vol. 154, pp. 219-233, 2016a.

[3] C. Malings and M. Pozzi, "Conditional entropy and value of information metrics for optimal sensing in infrastructure systems," *Structural Safety,* vol. 60, pp. 77-90, 2016b.

[4] C. Malings and M. Pozzi, "Value-of-information in spatio-temporal systems: Sensor placement and scheduling," *Reliability Engineering & System Safety,* vol. 172, pp. 45-57, 2018.

[5] F. A. Khan, S. A. Khan, D. Turgut and L. Bölöni, "Greedy path planning for maximizing value of information in underwater sensor networks," in *39th Annual IEEE Conference on Local Computer Networks Workshops*, Edmonton, 2014.

[6] L. Bölöni, D. Turgut, S. Basagni and C. Petrioli, "Scheduling data transmissions of underwater sensor nodes for maximizing value of information," in *IEEE Global Communications Conference (GLOBECOM),* Atlanta, GA, 2013.

[7] P. Gjanci, C. Petrioli, S. Basagni, C. A. Phillips, L. Bölöni and D. Turgut, "Path Finding for Maximum Value of Information in Multi-Modal Underwater Wireless Sensor Networks," *IEEE Transactions on Mobile Computing,* vol. 17, no. 2, pp. 404-418, 2018.

[8] M. H. Faber, Risk and safety in civil engineering, Zurich: Swiss Federal Institute of Technology ETH, 2007.

[9] D. Straub and A. Der Kiureghian, "Bayesian network enhanced with structural reliability methods: Application," *Journal of Engineering Mechanics,* vol. 136, no. 10, pp. 1259-1270, 2010.

[10] C. Bisdikian, L. M. Kaplan and M. B. Srivastava, "On the quality and value of information in sensor networks," *ACM Transactions on Sensor Networks (TOSN),* vol. 9, no. 4, p. 48, 2013.

[11] C. Malings and M. Pozzi, "Submodularity issues in value-of-information-based sensor placement," *Reliability Engineering & System Safety* , vol. 183 , pp. 93-103, 2019.

[12] A. Krause, A. Singh and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *Journal of Machine Learning Research,* vol. 9, pp. 235-284, 2008.

[13] C. Guestrin, A. Krause and A. P. Singh, "Near-optimal sensor placements in gaussian processes," in *22nd international conference on Machine learning,* ACM, 2005.

[14] A. Krause, H. B. McMahan, C. Guestrin and A. Gupta, "Robust submodular observation selection," *Journal of Machine Learning Research,* vol. 9(Dec), pp. 2761-2801, 2008.

[15] H. Raiffa and R. Schlaifer, Applied Statistical Decision Theory, Harvard University, 1961.

[16] R. A. Howard, "Information value theory," *IEEE Transactions on systems science and cybernetics,* vol. 2, no. 1, pp. 22-26, 1966.

[17] G. L. Nemhauser, L. A. Wolsey and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—I," *Mathematical programming* , vol. 14, no. 1, pp. 265-294, 1978.

[18] DI MAIO, F., Hoseyni, S. M., & Zio, E., "Stima adattativa del rischio di rottura di componenti in pressione soggetti a creep con un approccio probabilistico". In *SAFAP 2018-Sicurezza ed affidabilità delle attrezzature a pressione,* pp. 1-9, Bologna, 2018.

Authors' background (This form is only for submitted manuscript for review)

Your Name	Title*	Affiliation	Research Field	Personal website
Seyed Mojtaba Hoseyni	PhD candidate	Energy Department, Politecnico di Milano, Milan, Italy	Probabilistic Safety Assessment of complex systems, Maintenance, Decision Analysis.	www.lasar.polimi.it
Francesco Di Maio	associate professor	Energy Department, Politecnico di Milano, Milan, Italy	Safety analysis of complex systems by soft computing techniques (neural networks, fuzzy logic, genetic algorithms).	www.lasar.polimi.it
Enrico Zio	full professor	[1]Energy Department, Politecnico di Milano, Milan, Italy [2]Centre de Recherche sur les Risques et les Crises (CRC), MINES ParisTech/PSL Université Paris, Sophia Antipolis, France [3]Department of Nuclear Engineering, Kyung Hee University, Seoul, South Korea	Risk Analysis, Reliability, resilience, optimization, prognostics and health management	www.lasar.polimi.it

*This form helps us to understand your paper better, the form itself will not be published.

*Title can be chosen from: master student, Phd candidate, assistant professor, lecture, senior lecture, associate professor, full professor