

Modelo para imputar valores perdidos en evaluación de *Nuclear Safeguards*

R.M. Rodríguez

Dept. de Informática

Universidad de Jaén

rmrodrig@ujaen.es

D. Ruan

SCK • CEN

Nuclear Research Centre

druan@sckcen.be

J. Liu

School of Comp. and Maths

University of Ulster

j.liu@ulster.ac.uk

L. Martínez

Dept. de Informática

Universidad de Jaén

martin@ujaen.es

Resumen

La evaluación de *nuclear safeguards* tiene como objetivo verificar que los países no están desarrollando armas o actividades nucleares. Los inspectores de la Agencia de Energía Atómica Internacional (IAEA) evalúan indicadores utilizando diferentes fuentes de información que en ocasiones son vagas e imprecisas. El uso de la información lingüística ha proporcionado una forma de manejar tal incertidumbre. Sin embargo, pueden aparecer valores perdidos en las evaluaciones proporcionadas por los inspectores, ya que hay más de 900 indicadores. Ésto puede ser por la falta de experiencia o conocimiento de los inspectores sobre algunos indicadores. Los valores perdidos pueden realizar una desviación en el resultado final del proceso de evaluación. En esta contribución proponemos un proceso de imputación basado en filtrado colaborativo. Además, presentamos una medida de confianza que indica la bondad de las estimaciones obtenidas para los valores perdidos.

1. Introducción

Nuclear safeguards consiste en un conjunto de actividades de las que se encarga la Agencia de Energía Atómica Internacional (IAEA), para verificar que un país no está desarrollando armas o actividades nucleares. El control de *nuclear safeguards* se basa en la evaluación que realiza la IAEA para detectar tales actividades, utilizando para ello diferentes fuentes de información.

Para llevar a cabo dicha evaluación, la IAEA ha desarrollado un Modelo Físico [6] basado en una estructura jerárquica que permite incluir todas las actividades que podrían estar implicadas en el desarrollo de armas nucleares, que van desde indicadores básicos a factores complejos, sintetizados a partir de los indicadores, que indican el desarrollo de actividades nucleares (ver Fig. 1).

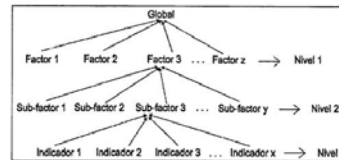


Figura 1: Estructura de la evaluación global

Los inspectores de la IAEA evalúan los indicadores basándose en sus análisis y conocimientos según las diferentes fuentes de información disponibles que suelen ser vagas e imprecisas [11, 12]. Inicialmente el principal problema en *nuclear safeguards* era el tratamiento de dicha incertidumbre que ha sido satisfactoriamente resuelto en [11, 12, 13] mediante modelos de evaluación, basados en el uso del enfoque lingüístico difuso [21]. Por esta razón, el foco de atención en [11, 12, 13] se ha desplazado a mejorar los resultados cuando aparecen valores perdidos en la evaluación. Ésta es una situación habitual por la complejidad, la gran cantidad de indicadores y la imprecisión de la

información que forma parte del problema.

El tratamiento de valores perdidos se ha realizado con diversas técnicas [8, 15, 16]: *eliminación, imputación y usarlos tal como son*. Es necesario estudiar qué estrategia es mejor en cada problema y situación. En *nuclear safeguards* se ha propuesto un modelo basado en la estrategia *usarlos tal como son* [8]. Sin embargo, en esta contribución se propone un modelo de *imputación* basado en filtrado colaborativo (FC) con un esquema Knn [2, 3, 4]. Un sistema de FC con un esquema Knn es un sistema que estima valoraciones basándose solamente en los términos de similitud (K similitudes más cercanas) entre inspectores. Una imputación consiste en sustituir un valor perdido por una estimación que representaría el valor que el inspector hubiese proporcionado si hubiese tenido el conocimiento o información suficiente para realizarlo.

La razón de usar un modelo basado en FC para imputar valores perdidos en *nuclear safeguards*, es porque el conjunto de datos utilizado en dicha evaluación se ajusta a los requerimientos del sistema de FC. Aun así, es necesario realizar un estudio sobre cada conjunto de datos para optimizar el modelo de imputación.

Dado que los valores imputados no son valores reales sino estimaciones, es importante conocer la fiabilidad de los mismos. Por lo que también se propondrá una medida de confianza para calcular dicha fiabilidad.

El resto del trabajo es estructurado como sigue: En la Sección 2 se revisan algunos trabajos acerca de *nuclear safeguards*. En la Sección 3 se introducen preliminares necesarios. En la Sección 4, se presentan una nueva propuesta para manejar valores perdidos en *nuclear safeguards* mediante un modelo de imputación y una medida de confianza. En la Sección 5, se muestra un caso de estudio de la metodología propuesta, y finalmente la Sección 6 concluye esta contribución.

2. Trabajos relacionados

El modelo físico que utiliza la IAEA proporciona una estructura para organizar la informa-

ción en *nuclear safeguards*. Esta información es utilizada por los inspectores de la IAEA para detectar la posible existencia de actividad nuclear en un país. Dicho modelo define indicadores para un determinado proceso que existe o está en desarrollo. La IAEA utiliza 914 indicadores que se clasifican según su importancia en fuertes, medios y débiles.

En [11] se presentó un modelo de evaluación lingüístico para el tratamiento de *nuclear safeguards* basado en el modelo simbólico computacional 2-tupla [5]. Este modelo usa el modelo físico de la IAEA y está dividido en varios niveles con diferente complejidad. La valoración global se obtiene mediante un proceso de agregación multi-nivel.

En [12] fue presentado el modelado, análisis y agregación de la información en *nuclear safeguards* bajo diferentes tipos de incertidumbre. Este enfoque utiliza el modelo de evaluación multi-nivel presentado en [11], y un nuevo modelo de inferencia basado en la metodología (RIMER) [20] para tratar información imprecisa en *nuclear safeguards*.

3. Preliminares

La evaluación de *nuclear safeguards* trata con información imprecisa relacionada con procesos cognitivos. El uso del enfoque lingüístico ha proporcionado buenos resultados para manejar este tipo de información en procesos nucleares [11, 12]. Dado que en nuestra propuesta de imputación de valores perdidos utilizamos estos modelos lingüísticos que implican a su vez procesos de computación con palabras (CP), en esta sección revisaremos brevemente el enfoque lingüístico y algunos conceptos de CP.

3.1. Enfoque lingüístico

El enfoque lingüístico difuso [21] representa la información mediante descriptores lingüísticos y su semántica asociada. Una forma de elegir los descriptores lingüísticos puede ser proporcionando directamente el conjunto de términos, distribuidos en una escala que tiene un orden definido [19]. Por ejemplo, un conjunto

de 7 términos lingüísticos podría ser:

$$S = \{s_0 : nada(n), s_1 : muy\ bajo(mb), s_2 : bajo(b), s_3 : medio(m), s_4 : alto(a), s_5 : muy\ alto(ma), s_6 : perfecto(p)\}$$

En este caso, se requiere que existan los operadores: i) Negación: $Neg(s_i) = s_j$ tal que $j = g - i$ ($g + 1$ es la cardinalidad), (ii) Maximización: $máx(s_i, s_j) = s_i$ si $s_i \geq s_j$, (iii) Minimización: $mín(s_i, s_j) = s_i$ si $s_i \leq s_j$.

La semántica de los términos se da mediante números difusos. Una forma de caracterizar un número difuso es utilizar una representación paramétrica de sus funciones de pertenencia.

3.2. Modelo lingüístico computacional 2-tupla

El modelo lingüístico computacional 2-tupla fue introducido en [5] para mejorar la precisión en los procesos de CP, y en [11, 12] fue utilizado para modelar la incertidumbre en la evaluación de la protección nuclear y realizar los procesos computacionales. Este modelo representa la información lingüística mediante un par de valores, denominado 2-tupla, (s_i, α) , donde s_i es un término lingüístico y α es un valor numérico que representa la traslación simbólica.

Definición 1 La traslación simbólica de un término lingüístico $s_i \in S = \{s_0, \dots, s_g\}$ es un valor numérico definido en $[-0.5, 0.5]$ que representa la "diferencia de información" entre una cantidad de información $\beta \in [0, g]$ obtenida de una operación simbólica y el índice del término lingüístico más cercano.

Este modelo define un conjunto de funciones para realizar transformaciones entre valores numéricos y 2-tupla para facilitar los procesos de CP.

Definición 2 Sea $S = \{s_0, \dots, s_g\}$ un conjunto de términos lingüísticos. La 2-tupla asociada a S es definida como $\langle S \rangle = S \times [-0.5, 0.5]$. La función $\Delta : [0, g] \rightarrow \langle S \rangle$ es definida mediante

$$\Delta(\beta) = (s_i, \alpha), \begin{cases} i = \text{round}(\beta), \\ \alpha = \beta - i, \end{cases} \quad (1)$$

donde round es el operador de redondeo, s_i es la etiqueta con índice más cercano a β y α es el valor de la traslación simbólica.

Δ es biyectiva y $\Delta^{-1} : \langle S \rangle \rightarrow [0, g]$ es definida mediante $\Delta^{-1}(s_i, \alpha) = i + \alpha$. De esta forma la 2-tupla de $\langle S \rangle$ es identificada con el valor numérico en el intervalo $[0, g]$.

Además, junto a este modelo de representación, un modelo computacional basado en las funciones Δ y Δ^{-1} fue también definido en [5].

4. Modelo para imputar valores perdidos en nuclear safeguards

Hasta hace poco tiempo, el interés en la evaluación de *nuclear safeguards* se centraba en el desarrollo del modelo de evaluación cuya estructura general se muestra en la Fig. 1. Los indicadores y sub-factores son agregados por niveles hasta obtener una valoración global. Sin embargo, después de los modelos presentados en [11, 12], el interés en este problema se ha focalizado en el tratamiento de valores perdidos [8, 9], ya que el tratamiento de dichos valores es clave para obtener resultados más fiables.

De las distintas estrategias de tratamiento de valores perdidos: *eliminación, imputación y usarlos tal como son* [7, 15, 17], esta contribución se centra en la estrategia de imputación [14, 16, 18], que presenta un proceso de imputación para *nuclear safeguards* basado en FC y una medida de confianza para obtener la fiabilidad de los valores imputados. Por tanto, el proceso de evaluación se realiza tal y como se indica en la Fig. 2. Se obtienen las valoraciones de los inspectores, se estiman los valores perdidos y se agregan dichas valoraciones para obtener una valoración global.

4.1. Proceso de imputación

El proceso de imputación está basado en un esquema Knn y un proceso de imputación similar a un sistema de recomendación colaborativo [1, 3] (ver Fig. 3).

Primero agrupamos los indicadores según su similitud. Para ello utilizamos una técnica de FC basada en el esquema Knn (K vecinos más cercanos) que es uno de los más usados por

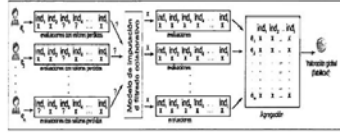


Figura 2: Pasos generales para evaluación en protección nuclear con valores perdidos

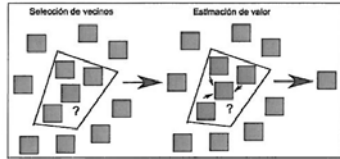


Figura 3: Esquema de un sistema de recomendación colaborativo

su robustez y precisión. El esquema Knn se ha aplicado con éxito en diferentes campos [2, 3, 4]. Dado que no existe una configuración óptima para un conjunto de datos, es necesario realizar un estudio previo para ajustar diferentes parámetros y medidas que proporcionen agrupaciones útiles. En la Sección 5 mostraremos un caso de estudio.

Para realizar las agrupaciones debemos elegir una medida de similitud. En [3] fueron introducidas dos medidas de similitud como las más utilizadas en FC. En nuestra propuesta hemos modificado las definiciones originales para tratar con 2-tuplas lingüísticas [5].

- Medida de similitud del coseno:

$$w(l, m) = \frac{\sum_k \Delta^{-1}(x_l^k) \Delta^{-1}(x_m^k)}{\sqrt{\sum_k \Delta^{-1}(x_l^k)} \sqrt{\sum_k \Delta^{-1}(x_m^k)}} \quad (2)$$

donde l y m son indicadores y x_l^k es la valoración proporcionada por el inspector e_k al indicador l .

- Coeficiente de correlación de Pearson:

$$w(l, m) = \frac{(\sum_k \Delta^{-1}(x_l^k) - \Delta^{-1}(\bar{x}_l)) \times (\sum_k \Delta^{-1}(x_m^k) - \Delta^{-1}(\bar{x}_m))}{\sqrt{\sum_k (\Delta^{-1}(x_l^k) - \Delta^{-1}(\bar{x}_l))^2} \times \sqrt{\sum_k (\Delta^{-1}(x_m^k) - \Delta^{-1}(\bar{x}_m))^2}} \quad (3)$$

siendo x_l^k la valoración proporcionada por el inspector e_k al indicador l , y \bar{x}_l la valoración media del indicador l .

Una vez calculadas las agrupaciones, el modelo de FC estima un valor para aquel indicador que aún no ha sido valorado.

Las estimaciones se calculan mediante un proceso de agregación de los K vecinos más cercanos seleccionados previamente. Existen diferentes métodos de agregación, aquí hemos estudiado los dos más habituales en FC:

- Suma ponderada

$$x_l^e = \Delta \left(\frac{\sum_{m=1}^{m=k} w(l, m) \Delta^{-1}(x_m^e)}{\sum_{m=1}^{m=k} w(l, m)} \right) \quad (4)$$

- Media + ajuste

$$x_l^e = \Delta \left(\frac{\sum_{m=1}^{m=k} w(l, m) \Delta^{-1}(x_m^e) - \bar{x}^e}{\sum_{m=1}^{m=k} |w(l, m)|} \right) \quad (5)$$

donde l es un indicador, $e \in E$, es el inspector que no ha proporcionado su valoración, y k , es el número de indicadores seleccionados para calcular el valor, x_l^e .

Resaltar que utilizamos las funciones Δ y Δ^{-1} para transformar entre valores numéricos y 2-tupla.

4.2. Fiabilidad de los valores imputados

Los métodos de FC proporcionan métricas como MAE (Error Absoluto Medio)[4], ROC [10], etc., para calcular el error cometido en las estimaciones. Sin embargo estas métricas no son suficientes, ya que sólo indican un error medio y debido a las necesidades de *nuclear safeguards*, sería necesario un valor que indique

Cuadro 1: Configuraciones y resultados experimentales

	Cos y Suma pond.			Cos y Media + ajuste		
	K=15	K=10	K=5	All-1	n=8	n=4
	Cf_1	Cf_2	Cf_3	Cf_4	Cf_5	Cf_6
MAE (10 %)	1.17	1.23	1.31	1.21	1.23	1.23
MAE (20 %)	1.23	1.27	1.37	1.27	1.30	1.35
	Pearson y Suma pond.			Pearson y Media + ajuste		
	K=15	K=10	K=5	All-1	n=8	n=4
	Cf_7	Cf_8	Cf_9	Cf_{10}	Cf_{11}	Cf_{12}
MAE (10 %)	1.48	1.50	1.59	1.49	1.51	1.57
MAE (20 %)	1.74	1.8	1.87	1.78	1.81	1.89

la confianza de los valores imputados para calcular la fiabilidad del resultado. Para alcanzar este objetivo se define una medida que proporciona la confianza de los valores imputados. Esta medida se basa en diversos casos de estudio (Sección 5, [2]) y se define como:

$$T(\bar{x}_i^l) = (1 - \overline{sim}_l)h + \overline{sim}_l \frac{k}{K}, \quad T(\bar{x}_i^l) \in [0, 1] \quad (6)$$

$$h = \frac{g - sd(\bar{x}_i^l)}{g}$$

donde \overline{sim}_l es la media aritmética de las similitudes entre el indicador, l , y sus k vecinos más cercanos. Y h muestra la homogeneidad de las valoraciones x_i^l , siendo sd la desviación estándar de las valoraciones usadas para calcular el valor imputado, y $g+1$ la granularidad de S . Finalmente, k indica el número real de vecinos implicados en el cálculo de \bar{x}_i^l de los iniciales K calculados con el esquema Knn.

Esta medida indica que cuantas más valoraciones utilizamos para realizar las estimaciones, mayor es la fiabilidad. De igual forma, cuanto más homogéneas sean las valoraciones, más fiables son los valores imputados. Por tanto, cuanto mayor sea $T(\bar{x}_i^l)$, más fiable será el valor.

Esta confianza, $T(\Delta)$, es usada en *nuclear safeguards*, realizando un proceso de agregación similar al realizado con los indicadores, para medir la confianza del resultado final.

5. Caso de estudio

En esta sección presentamos un ejemplo de aplicación de la propuesta anterior. Este ejemplo usa un conjunto de datos reducido de *nuclear safeguards*, el cual ha sido también utilizado en otros modelos [8, 9].

Teniendo en cuenta que no existe un modelo de FC óptimo basado en el esquema Knn para cualquier problema y conjunto de datos, presentamos el caso de estudio que hemos realizado para ajustar el modelo de FC, y a continuación mostramos el proceso de imputación.

5.1. Ajuste de parámetros para el conjunto de datos de *nuclear safeguards*

En este caso de estudio utilizamos un conjunto de datos reducido de *nuclear safeguards* que consta de 22 indicadores de los 914 usados en el proceso completo, en el que 4 inspectores expresan sus valoraciones utilizando la sintaxis descrita en la Sección 3.1.

El objetivo de este caso de estudio es optimizar los diferentes parámetros del modelo de FC basado en el esquema Knn, para obtener buenas estimaciones. Los parámetros son: (i) Tamaño del vecindario, (ii) Medida de similitud y (iii) Método de estimación.

Cada configuración, Cf_i , del Cuadro 1 se ha ejecutado 500 veces, K indica el número de vecinos de las agrupaciones utilizadas. En cada configuración hemos utilizado aproximadamente el 10% y 20% de valores perdidos de las valoraciones proporcionadas por los inspectores, y hemos usado el MAE para medir

Cuadro 2: Evaluaciones de los expertos

ind.	1	2	3	4	5	6	7	8	9	10	11
p1	a	p	ma	m	?	a	m	p	m	b	?
p2	b	ma	?	m	b	m	b	mb	b	b	m
p3	a	a	p	ma	m	ma	ma	a	?	m	ma
p4	p	p	p	?	p	m	ma	ma	a	a	m
ind.	12	13	14	15	16	17	18	19	20	21	22
p1	mb	p	ma	?	b	a	m	m	p	a	m
p2	m	p	m	b	b	m	m	?	m	m	mb
p3	b	p	p	mb	b	p	m	ma	p	ma	?
p4	?	ma	a	b	m	?	b	ma	a	p	mb

Cuadro 3: Estimaciones y fiabilidades

ind.	e1		e2		e3		e4	
	estim.	fiab.	estim.	fiab.	estim.	fiab.	estim.	fiab.
3			(m,-.154)	.866				
4							(a,.313)	.866
5	(a,-.46)	.863						
9					(ma,-.361)	.931		
11	(a,-.132)	.93						
12							(a,.242)	.796
15	(m,.493)	.798						
17							(ma,-.459)	.866
19			(m,-.16)	.866				
22					(ma,-.4)	.858		

su precisión. La mejor configuración obtenida para el proceso de imputación del ejemplo que vamos a desarrollar ha sido Cf_1 .

5.2. Proceso de imputación basado en el FC para nuclear safeguards. Ejemplo ilustrativo

Aquí mostramos la funcionalidad del proceso de imputación en nuclear safeguards. Para ello utilizamos el conjunto de datos del Cuadro 2, donde 4 inspectores expresan sus valoraciones sobre 22 indicadores. El número de valores perdidos es 10.

Utilizando la configuración Cf_1 , obtenemos los valores imputados y la confianza mostrada en el Cuadro 3.

A continuación detallamos el cálculo del valor perdido, x_5^1 , proporcionado por el inspector, e_1 , en el indicador, ind_5 .

- Primero calculamos la similitud entre el ind_5 y los restantes indicadores, utilizando la Eq. 2 y $K=15$:

$$w(5, 1) = \frac{(4 \cdot null) + (2 \cdot 2) + (4 \cdot 3) + (6 \cdot 6)}{\sqrt{null+4+16+36} \sqrt{null+4+9+36}} = \frac{4+12+36}{\sqrt{56} \sqrt{49}} = 0.993$$

- La estimación del valor imputado mediante la Eq. 4 es:

$$x_5^1 = x_5^1 = \Delta(3.54) = (alto, -.46)$$

- Una vez calculada la estimación del valor perdido, el modelo calcula su confianza según la Eq. 6. Para utilizar dicha función es necesario calcular los siguientes parámetros:

$$sd = 1.45, \quad g = 6, \quad h = 0.759, \quad \overline{sim}_5 = 0.968 \quad y \quad k = 13$$

Por tanto, la confianza del valor imputado x_5^1 es:

$$T(x_{15}^1) = (1 - 0.968) * 0.759 + 0.968 * 0.867 = 0.863$$

El valor imputado para el valor perdido del indicador 5 proporcionado por el inspector 1 es (alto, -.46) y su confianza 0.863.

Tras la imputación de todos los valores perdidos se realiza el proceso de agregación multinivel presentado en la Fig. 2 y obtenemos una valoración global y su fiabilidad.

Existen diferentes operadores de agregación. Nosotros por simplicidad utilizamos el operador de la media con el que obtenemos los siguientes resultados:

- Valoración global para *nuclear safeguards*: (alto, -.23)
- Fiabilidad: 0.864

Debemos mencionar que si utilizamos el modelo de *nuclear safeguards* presentado en la Fig. 1 sin realizar el proceso de imputación, el resultado es: (medio, .33).

Por tanto podemos observar la importancia del tratamiento de los valores perdidos en *nuclear safeguards*, ya que su tratamiento puede obtener diferentes resultados.

6. Conclusión

La evaluación de *nuclear safeguards* es un proceso complejo en el que hay que valorar varios cientos de indicadores y la información que participa en el mismo suele ser vaga e imprecisa. Tras la definición de modelos para manejar dicha incertidumbre y complejidad, uno de los problemas más importantes a los que se enfrenta la IAEA es cómo tratar los valores perdidos que aparecen en las evaluaciones de sus inspectores.

En esta contribución se ha propuesto un modelo de imputación de valores perdidos basado en filtrado colaborativo para obtener los valores a imputar. También se ha presentado una medida de fiabilidad de los mismos, para saber la confianza que se puede tener en la evaluación después de la imputación de los valores calculados.

De los resultados obtenidos, parece interesante cara al futuro, estudiar modelos híbridos

de tratamiento de valores perdidos que dependerán del conocimiento que se tenga en cada evaluación.

Agradecimientos

Este trabajo está parcialmente subvencionado por los proyectos de investigación TIN-2009-08286, P08-TIC-3548 y fondos FEDER.

Referencias

- [1] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Ieee Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005.
- [2] E.J. Castellano and L. Martínez. A web-decision support system based on collaborative filtering for academic orientation. case study of the spanish secondary school. *Journal of Universal Computer Science*, 15(14):2786–2807, 2009.
- [3] J.L. Herlocker, J.A. Konstan, and J. Riedl. An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms. *Information Retrieval*, 5(4):287–310, 2002.
- [4] L.J. Herlocker, J.A. Konstan, L.G. Terveen, and J.T. Riedl. Evaluating collaborative filtering recommender systems. In *ACM Transactions on Information Systems*, volume 22, pages 5–53, 2004.
- [5] F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Trans. on Fuzzy Systems*, 8(6):746–752, 2000.
- [6] IAEA. *IAEA Bulletin, Annual Report*, volume 43, chapter Nuclear Security and Safeguards. 2001.
- [7] N. Jiang and L. Gruenwald. *Advances in Databases: Concepts, Systems and Applications*, volume 4443/2008, chapter Estimating missing data in data streams, pages 981–987. R. Kotagiri and P.R. Krish-

- na and M. Mohania and E. Nantajeewarawat, 2007.
- [8] Ö. Kabak and D. Ruan. A cumulative belief degree-based approach for missing values in nuclear safeguards evaluation. *IEEE Transactions on Knowledge and Data Engineering*, page (<http://doi.ieeecomputersociety.org/10.1109/TKDE.2010.60>), 2010.
- [9] Ö Kabak and D. Ruan. A cumulative belief-degree approach for nuclear safeguards evaluation. In *Proceedings of the 2009 IEEE International Conference on Systems, Man and Cybernetics*, pages 2285–2290, San Antonio, TX, USA, 2009.
- [10] T.C.W. Landgrebe and R.P.W. Duin. Efficient multiclass ROC approximation by decomposition via confusion matrix perturbation analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5):810–822, 2008.
- [11] J. Liu, D. Ruan, and R. Carchon. Synthesis and evaluation analysis of the indicator information in nuclear safeguards applications by computing with words. *International Journal of Applied Mathematics and Computer Science*, 12(3):449–462, 2002.
- [12] J. Liu, D. Ruan, H. Wang, and L. Martínez. Improving nuclear safeguards evaluation through enhanced belief rule-based inference methodology. *Int. J. Nuclear Knowledge Management*, 3(3):312–339, 2009.
- [13] L. Maschio. A decision support system for safeguards information analysis. *International Journal of Nuclear Knowledge Management*, 2(4):410–421, 2007.
- [14] R. Nowicki. On combining neuro-fuzzy architectures with the rough set theory to solve classification problems with incomplete data. *IEEE Transactions on Knowledge and Data Engineering*, 20(9):1239–1253, 1989.
- [15] L.B. Oltman and S.B. Yahia. *Concept Lattices and their applications*, volume 4923/2008, chapter Yet another approach for completing missing values, pages 155–169. S.B. Yahia and E.M. Nguifo and R. Belohlavek, 2008.
- [16] M. Pawlak. Kernel classification rules from missing data. *IEEE Transactions on Information Theory*, 39(3):979–988, 1993.
- [17] J. Siddique and R. Belin. Using and approximate bayesian bootstrap to multiply impute nonignorable. *Comp. Statistics and Data Analysis*, 53(2):405–415, 2008.
- [18] R. Slowinski and J. Stefanowski. Rough classification in incomplete information-systems. *Mathematical and computer modelling*, 12(10-11):1347–1357, 1989.
- [19] R.R. Yager. An approach to ordinal decision making. *International Journal of Approximate Reasoning*, 12:237–261, 1995.
- [20] J. B. Yang, J. Liu, J. Wang, and H. S. Sii. A belief rule base inference methodology using the evidential reasoning approach - RIMER. *IEEE Transactions on Systems, Man, and Cybernetics*, 36(2):266–285, 2006.
- [21] L.A. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning. *Inf. Sci., Part I, II, III*, 8,8,9:199–249,301–357,43–80, 1975.