Revista Cubana de Ciencias Informáticas
Vol. 11, No. 4, Octubre-Diciembre, 2017
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 57- 66

Tipo de artículo: Artículo original
Temática: Inteligencia Artificial
Recibido: 11/02/2017 | Aceptado: 23/09/2017

El riesgo de seguridad de la información en gestores de bases de datos basado en números difusos trapezoidales

The security risk of information in database managers based on trapezoidal fuzzy numbers

Yasser Azán-Basallo[1], Natalia Martínez Sánchez[2], Vivian Estrada Senti[3]

[1]Centro Telemática, Facultad 2, Universidad de las Ciencias Informáticas. Carretera a San Antonio de los Baños km 2½, Reparto. Torrens. Boyeros, La Habana. CUBA. Correo electrónico: yazan@uci.cu

[2] Vicerrectoría de Formación, Universidad de las Ciencias Informáticas. Carretera a San Antonio de los Baños km 2½, Reparto. Torrens. Boyeros, La Habana. CUBA. Correo electrónico: natalia@uci.cu

[3]Centro Internacional del Postgrado, Universidad de las Ciencias Informáticas. Carretera a San Antonio de los Baños km 2½, Reparto. Torrens. Boyeros, La Habana. CUBA. Correo electrónico: vivian@uci.cu

[*] Autor para correspondencia: yazan@uci.cu

Resumen

En este trabajo se plantea una solución para evaluar el riesgo de la seguridad de la información para los gestores de bases de datos. Teniendo como premisa de entrada, el riesgo cualitativo de cada parámetro de la lista de chequeo de seguridad. Para lograr este objetivo se propone la utilización de números difusos trapezoidales basado en el área. Se realiza una revisión de la literatura de nuevas funciones de semejanza, las cuales utilizan los números difusos, para argumentar la selección del modelo escogido para resolver el objetivo en este trabajo. Se muestra un caso de estudio para explicar el funcionamiento del modelo propuesto e indicar la viabilidad del mismo para manejar los términos lingüísticos de las premisas de entrada y lograr la determinación de la evaluación del riesgo de seguridad de la información igualmente de una manera cualitativa. Además, mostrar los resultados obtenidos con la propuesta a través de una herramienta desarrollada que permitió automatizar el modelo propuesto.

Palabras clave: auditoría, evaluación, lógica difusa, números difusos trapezoidales, riesgo

Revista Cubana de Ciencias Informáticas
Vol. 11, No. 4, Octubre-Diciembre, 2017
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 57- 66

Abstract

In this paper is proposed one solution for assessing the information security risk for database managers. It´s based on the qualitative risk of each parameter in the security checklist. To achieve this objective is proposing the use of trapezoidal fuzzy numbers based on area. A review of the literature of new similarity functions, which use the fuzzy numbers, is made to argue the selection of the chosen model to solve the objective in this work. A case study is presented to explain the operation of the proposed model and indicate its feasibility to handle the linguistic terms of the input premises and to achieve the determination of the information security risk assessment in a qualitative way. Also show the results obtained with the proposal through a tool developed that allowed to automate the proposed model.

Keywords: *audit, evaluation, diffuse logic, diffuse trapezoidal numbers, risk*

Introducción

En una anterior publicación se ha abordado la evaluación del riesgo de la seguridad de la información en los sistemas gestores de bases de datos (SGBD) utilizando el razonamiento basado en casos (Azán , Y. B. y otros, 2014). En esta se proporciona una evaluación del riesgo de forma cualitativa. El cual se realizó a partir de la obtención del riesgo cuantitativo de cada parámetro de la lista de chequeo de seguridad (Broder, J. F. y Tucker, G., 2011) que utilizan los expertos en auditoría de seguridad informática para los SGBD. También se publicó otro trabajo con el mismo objetivo pero utilizando la lógica difusa (Azán, Y. B., Martínez, N. S., y Estrada, S. V., 2015). Estos trabajos muestran la viabilidad de ambas técnicas para la evaluación del riesgo de la seguridad de la información en los SGBD teniendo como premisas o entradas, valores numéricos. Sin embargo, no se ha encontrado una solución para llevar a cabo la evaluación del riesgo de seguridad de la información a partir de los valores cualitativos: Alto, Medio o Bajo y devolver estos mismos valores como resultado final.

Por tal motivo se trazó como objetivo, proponer la evaluación del riesgo de la seguridad de la información para los SGBD teniendo como premisa de entrada, el riesgo cualitativo de cada parámetro de la lista de chequeo de seguridad.

Revisión de la literatura

Se han desarrollado varias investigaciones con el objetivo de realizar la evaluación de riesgo a partir de valores cualitativos, aunque no afines a la seguridad de la información, utilizando la lógica difusa (Abbasianjahromi, H. y Rajaie, H., 2013; Patra, K. y Mondal, S. K., 2015; Schmucke, K. J., 1984; Wang, Y.-M. y Elhag, T. M., 2006; Xu, Z.,

Shang, S., Qian, W., y Shu, W., 2010). Se encuentran nuevas propuestas de funciones de similitud para números difusos. Para el autor de esta investigación resultan de interés los trabajos de Patra y Mondal (2015) y de Vicente, Mateos y Jiménez (2013) por los resultados evidenciados de superioridad de su propuesta ante otras existentes como se evidencia en el trabajo de los mismos para ser utilizado en esta investigación.

Además se encontró un trabajo el cual permite tomar decisiones a partir de factores de riesgo de varios proyectos de la construcción de Abbasianjahromi y Rajaie (2013) utilizando un modelo híbrido del RBC difuso con factores de riesgos cualitativos como rasgos predictores, para decidir cuál proyecto ejecutar. Este trabajo también puede ser otro acercamiento al problema planteado por esta investigación.

En esta investigación se decide utilizar la propuesta de Vicente, Mateos y Jiménez (2013) por la superioridad demostrada en el trabajo citado con los resultados obtenidas en las comparaciones realizadas por estos autores.

Materiales y métodos o Metodología computacional

Se necesita para aplicar la función de semejanza de Vicente, Mateos y Jiménez (2013) primeramente determinar el número difuso del rasgo objetivo. Esto se realizó a partir de la propuesta publicada en Patra y Mondal (2015) en la cual se sustituye la operación aritmética (impacto $\otimes$ Probabilidad de ocurrencia) por la variable riesgo local (RL) de cada parámetro de la lista de chequeo. La causa de esta modificación se debe a que se tiene como premisa el valor del RL, sin la necesidad de esta operación matemática, por lo que se simplifica el proceso de cómputo. Se incorpora una nueva variable (el peso) para ajustar el valor del RL, quedando como aparece en la ecuación 1:

$$R = (\sum_{i=1}^{n} peso_i \otimes RL_i) ø (\sum_{i=1}^{n} peso_i) \qquad (1)$$

En la anterior ecuación 1, la variable R representa el valor del número difuso del rasgo objetivo, que es a su vez el riesgo del servidor tal que R= (t1, t2, t3, t4; w). Son números reales t1, t2, t3, t4 y w tal que: $0 \leq t1 \leq t2 \leq t3 \leq t4 \leq 1$, $0 \leq w \leq 1$. La variable w representa la altura del trapecio.

Es importante utilizar en esta investigación el peso como medida de diferenciación entre los parámetros que componen la lista de chequeo de seguridad. Esto se debe a la existencia de parámetros que pueden tener igual nivel de RL, pero con contrastes más determinantes, en unos más que en otros, para la evaluación del riesgo del rasgo objetivo. Por tanto, la utilización del peso contribuye a la diferenciación y en la precisión de la estimación del valor de rasgo objetivo. El peso se determinó a través de encuestas aplicadas a los especialistas del Departamento de Seguridad Informática de la empresa ETECSA. El valor del peso está definido en el intervalo [0,1].

Revista Cubana de Ciencias Informáticas
Vol. 11, No. 4, Octubre-Diciembre, 2017
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 57- 66

El número difuso generalizado R, es un subconjunto borroso de la línea real $\mathbb{R}$, cuya función de pertenencia μ_R cumple las siguientes condiciones (Chen, S.-J. y Chen, S.-M., 2003):

1) μ_R es una correlación continua desde el intervalo cerrado [0,1];
2) $\mu_R(x) = 0$ donde $-\infty < x \leq a$;
3) $\mu_R(x)$ está aumentando estrictamente entre $[a, b]$;
4) $\mu_R(x) = w$ donde $b \leq x \leq c$;
5) $\mu_R(x)$ está disminuyendo estrictamente entre $[c, d]$;
6) $\mu_R(x) = 0$ donde $d \leq x < \infty$;

En la figura 1 se puede observar la representación gráfica de los números difusos trapezoidales generalizados de los valores lingüísticos creada con el MATLAB 7.6.0.

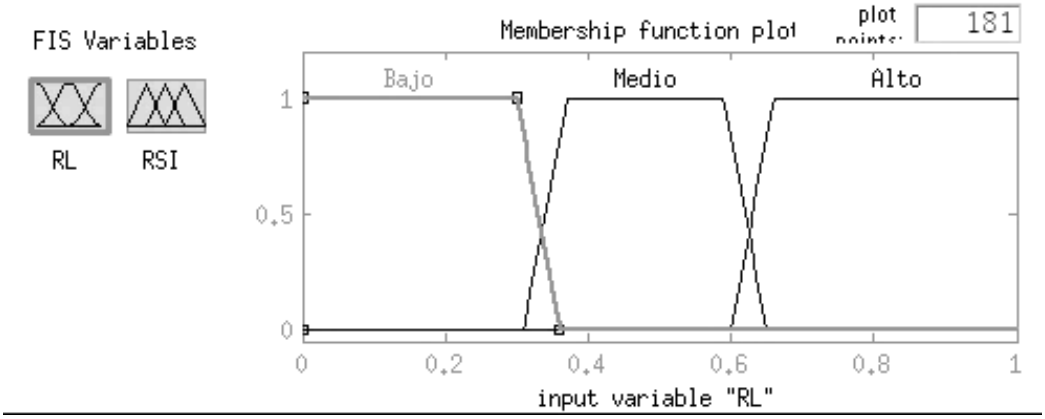

Figura 1. Representación gráfica de los números difusos trapezoidales generalizados de los valores lingüísticos

La función de pertenencia o miembro ($\mu(x)$: $\mathbb{R} \rightarrow [0, 1]$) trapezoidal generalizado es de la siguiente forma:

$$\mu_R(x) = \begin{cases} 0, & \text{si } x \leq t_1 \\ \frac{(x-t_1)}{t_2-t_1}, & \text{si } t_1 < x \leq t_2 \\ 1, & \text{si } t_2 < x \leq t_3 \\ \frac{(t_4-x)}{t_4-t_3}, & \text{si } t_3 < x \leq t_4 \\ 0, & \text{si } x > t_4 \end{cases} \quad (2)$$

Revista Cubana de Ciencias Informáticas
Vol. 11, No. 4, Octubre-Diciembre, 2017
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 57- 66

Función de semejanza

Para determinar el valor cualitativo del riesgo del rasgo objetivo, se utiliza una función de semejanza que permite comparar el riesgo del número difuso calculado por la ecuación (1) con los números difusos que están asociados a los valores lingüísticos de salida como se muestra en la tabla 1. A partir del valor de semejanza, tomar como resultado final el de mayor semejanza.

La propuesta general de función de semejanza seleccionada es la publicada por (Vicente, E. y otros, 2013):

Si máx $(|(X_{RN} - X_{RO})|, (|Y_{RN} - Y_{RO}|)) \neq 0$

$$S(RN, RO) = 1 - (1 - \alpha - \beta)\left(1 - \frac{\int_0^1 \mu_{RN \cap RO}(x)dx}{\int_0^1 \mu_{RN \cup RO}(x)dx}\right) - \alpha\frac{\sum|t_{RNi} - t_{ROi}|}{4} - \beta l_\infty[(X_{RN}, Y_{RN}), (X_{RO}, Y_{RO})] \quad (3)$$

En otro caso:

$$S(RN, RO) = 1 - \left(\frac{1-\alpha-\beta}{2} + \alpha\right)\frac{\sum|t_{RNi} - t_{ROi}|}{4} - \left(\frac{1-\alpha-\beta}{2} + \beta\right)|X_{RN} - X_{RO}| \quad (4)$$

Donde la variable RN es el riesgo expresado en un número difuso trapezoidal del nuevo caso de la auditoría de seguridad informática, la cual se desea diagnosticar o evaluar. La variable RO se corresponde a un número difuso de la tabla 1. El valor 1 representa la similitud exacta entre los casos, $\alpha + \beta < 1$. Las variables α y $\beta = {}^1/_3$ para que se puedan comparar los resultados con el análisis dispuesto en Vicente, Mateos y Jiménez (2013). El μ es la función miembro del número difuso R.

$$\beta l_\infty[(X_{RN}, Y_{RN}), (X_{RO}, Y_{RO})] = \underset{\alpha+\beta<1}{\overset{máx}{}}((|X_{RN} - X_{RO}|), (|Y_{RN} - Y_{RO}|)),\ \mu_{RN \cap RO}(x) =$$

$$\underset{[0\leq x\leq 1]}{\overset{mín}{}}(\mu_{RN}(x),\ \mu_{RO}(x)),\ \mu_{RN \cup RO}(x) = \underset{[0\leq x\leq 1]}{\overset{máx}{}}(\mu_{RN}(x), \mu_{RO}(x)) \quad (5)$$

(X_{RN}, Y_{RN}) y (X_{RO}, Y_{RO}) son los centroides de RN y RO y se calculan de la siguiente manera (Vicente, E. y otros, 2013) en la cual R es el número difuso trapezoidal como lo son RN y RO:

$$X_R = \left\{Y_{R(t_3+t2)+(w_R - Y_R)(t_4+t_1)}\right\},\ Y_R = \left\{\begin{array}{l} \frac{(\frac{t_3-t2}{t_4-t_1})}{6}, \text{si } t_4 - t_1 \neq 0 \\ \frac{1}{2},\ \text{si } t_4 - t_1 = 0 \end{array}\right\} \quad (6)$$

Revista Cubana de Ciencias Informáticas
Vol. 11, No. 4, Octubre-Diciembre, 2017
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 57- 66

Donde la variable RN es el riesgo expresado en un número difuso trapezoidal de la auditoría de seguridad informática, la cual se desea diagnosticar. La variable RO se corresponde al número difuso de la variable lingüística con la cual se quiere determinar el nivel de semejanza.

Tabla 1. Representación difusa de los términos lingüísticos.

Valores lingüísticos	Números difusos trapezoidales
Alto	(0.6, 0.66, 1, 1; 1)
Medio	(0.31, 0.37, 0.59, 0.65; 1)
Bajo	(0, 0, 0.3, 0.36; 1)

Caso de estudio

Para un mejor entendimiento de la propuesta de solución, los autores han preparado un ejemplo en un escenario real. En esta sección se ofrece en la tabla 2, un posible resultado de una auditoría a un SGBD en PostgreSQL en la cual va a tener los parámetros que fueron evaluados y el impacto correspondiente a cada uno de ellos.

Cuando se aplica la ecuación (1), se obtiene el número difuso de la evaluación del riesgo para esta auditoría R = (0.701697, 0.753032,0.884006, 0.930051,1).

El segundo paso, para obtener el diagnóstico de la evaluación de la auditoría, se emplea la ecuación (3). Obteniendo el nivel desemejanza entre la variable R con respecto a los números difusos de la tabla 1, los cuales representan los términos lingüísticos. El resultado de aplicar esta ecuación se muestra en la tabla 3, la cual muestra que el resultado del riesgo de la seguridad de la información para el SGBD PostgreSQL es de Alto.

Tabla 2. Ejemplo de evaluación de parámetros de una lista de chequeo de un SGBD PostgreSQL.

Nombre del parámetro	Impacto	Evaluación
Actualización del catálogo del sistema	**Alto**	**Mal**
Logines del motor de la base de datos	**Alto**	**Regular**
Pertenencia de usuarios a grupos.	**Medio**	**Bien**
Usuarios con claves nulas.	**Alto**	**Bien**
Cuentas vencidas	**Bajo**	**Bien**
Auditoría	**Alto**	**Regular**
Pertenencias de usuarios a grupos	**Medio**	**Bien**
Grupos vacíos	**Bajo**	**Bien**
Longines del motor de base de datos	**Regular**	**Alto**
Roles del motor de base de datos	**Alto**	**Regular**
Base de datos genéricas	**Medio**	Mal
Verificar la versión del motor de la base de datos	**Alto**	**Mal**

Revista Cubana de Ciencias Informáticas
Vol. 11, No. 4, Octubre-Diciembre, 2017
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 57- 66

Nombre del parámetro	Impacto	Evaluación
Ubicación de los ficheros log y los ficheros de datos	**Alto**	**Regular**
Modificaciones de la estructura de las tablas del sistema	**Alto**	**Bien**
Permisos de administración	**Alto**	**Regular**
Cantidad de conexiones concurrentes	**Bajo**	**Mal**
Permisos sobre la tabla pg_authid del catálogo del sistema	**Alto**	**Bien**
Parámetro listen_addreses	**Alto**	**Mal**
Parámetro password_encryption	**Alto**	**Bien**
Parámetro SSL	**Alto**	**Bien**
Parámetro unix_socket_permissions	**Alto**	**Regular**

Tabla 3. Resultado de la medición de la semejanza.

Alto	Medio	Bajo
0.9622	0.6491	0.3131

Resultados y discusión

Se desarrolló una solución informática a partir de la solución propuesta (Azán , Y. B. y otros, 2014) como instancia al modelo propuesto en este trabajo. La aplicación SASGBD como se le nombró, está creada a partir de su desarrollo anterior, por tanto, tiene las mismas facultades de diagnosticar los SGBD PostgreSQL, MySQL, Microsoft SQL Server y Oracle y además de mantener las mismas funcionalidades descritas en el anterior trabajo.

La misma es capaz de proponer el resultado del diagnóstico. El auditor es quien estima el resultado final de la auditoría de mantener la evaluación o de cambiarla.

Estudio comparativo

Se realizó un estudio comparativo para medir la exactitud de la propuesta de solución a través de casos de estudio proporcionados por los especialistas del Departamento de Seguridad Informática de la empresa ETECSA. Para evidenciar los resultados de la medición a través del SASGBD, los resultados fueron reflejados en la tabla 4. La columna R es de la palabra resultado, la variable C de correcto, M de Mal y A de aceptable.

Tabla 4. Resultado del valor de semejanza para el experimento uno.

Caso	Semejanza Alto	Semejanza Medio	Semejanza Bajo	Valor observado	Valor esperado	Resultado
O_1	0.028	0.339	0.964	Bajo	Bajo	Correcto
O_2	1.0	0.320	0.028	Alto	Alto	Correcto

Revista Cubana de Ciencias Informáticas
Vol. 11, No. 4, Octubre-Diciembre, 2017
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 57- 66

O_3	0.320	0.957	0.339	Medio	Medio	Correcto
O_4	0.363	0.883	0.310	Medio	Alto	Mal
O_5	0.257	0.823	0.428	Medio	Alto	Aceptable
O_6	0.208	0.694	0.517	Medio	Alto	Aceptable
O_7	0.206	0.688	0.521	Medio	Alto	Aceptable
O_8	0.252	0.811	0.436	Medio	Alto	Aceptable
O_9	0.115	0.481	0.729	Bajo	Bajo	Correcto
O_{10}	0.062	0.387	0.883	Bajo	Bajo	Correcto
O_{11}	0.066	0.388	0.835	Bajo	Bajo	Correcto
O_{12}	0.239	0.781	0.430	Medio	Alto	Mal
O_{13}	0.368	0.878	0.285	Medio	Alto	Mal
O_{14}	0.173	0.606	0.563	Medio	Alto	Mal
O_{15}	0.300	0.900	0.346	Medio	Alto	Mal

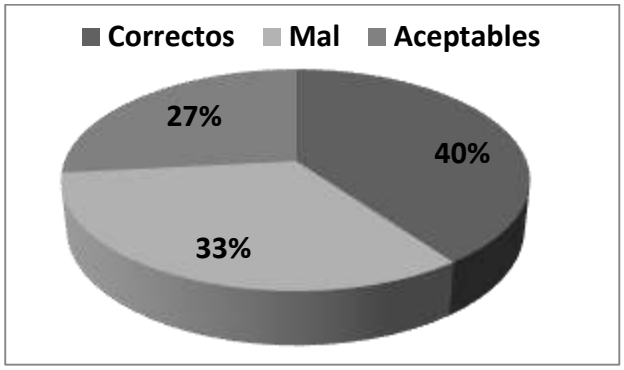

Figura 2. Resultados de la medición de la exactitud del primer experimento.

Los resultados de la medición de la exactitud del primer experimento mostrado en la figura 2 son que: de 15 casos, 6 fueron evaluados correctamente, 5 de mal y 4 de modo aceptable. Los casos de estudio valorados como correctos son aquellos donde coincide el valor esperado con el observado. Los aceptables son casos aquellos donde no coinciden los valores, pero el valor del observado es un resultado admisible. Los valorados de Mal son aquellos que los valores no coinciden y no son aceptados.

Conclusiones

La investigación realizada demuestra la viabilidad de realizar la evaluación del riesgo de seguridad de la información en gestores de bases de datos basado en la medida de la similitud de números difusos trapezoidales basado en el área.

Se evidencia a través de la propuesta de este trabajo, la oportunidad de manejar los términos lingüísticos que se ofrecen como premisa de entrada y de salida para la evaluación del riesgo de la seguridad de la información y mejorar la exactitud de la solución con el manejo de la ambigüedad existente en este proceso.

Referencias

ABBASIANJAHROMI, H. Y RAJAIE, H. (2013). Application of fuzzy CBR and MODM approaches in the project portfolio selection in construction companies. *Iranian Journal of Science and Technology. Transactions of Civil Engineering, 37,* 143-155. Extraído desde 20/02/2014, de http://www.sid.ir/En/VEWSSID/J_pdf/8542013C101.pdf.

AZÁN , Y. B., BRAVO , L. G., ROSALES , W. R., TRUJILLO , D. M., GARCÍA , E. A. R. Y PIMENTEL , A. R. (2014). Solución basada en el Razonamiento Basado en Casos para el apoyo a las auditorías informáticas a bases de datos. *Revista Cubana de Ciencias Informáticas, 8*, 52-68.

AZÁN, Y. B., MARTÍNEZ, N. S. Y ESTRADA, S. V. (2015). La lógica difusa para la evaluación del riesgo de seguridad informática a bases de batos. *Revista Control, Cibernética y Automatización, 3* 1-5. Extraído desde, de.

BRODER, J. F. Y TUCKER, G. (2011). Risk analysis and the security survey Extraído desde 12/12/2013, de http://www.google.com.cu/books?hl=es&lr=&id=fLmgIGT18jIC&oi=fnd&pg=PP1&dq=risk+assessment%2Bformula%2Bsecurity&ots=q1K-pmlGAY&sig=lwkAVW1G2jr7wkBjlR4dQftaQ2k&redir_esc=y#v=onepage&q=risk%20assessment%2Bformula%2Bsecurity&f=false.

CHEN, S.-J. Y CHEN, S.-M. (2003). Fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. *IEEE Transactions on fuzzy systems, 11*, 45-46. Extraído desde 14/6/2016, de http://ieeexplore.ieee.org/document/1178065/?denied.

PATRA, K. Y MONDAL, S. K. (2015). Fuzzy risk analysis using area and height based similarity measure on generalized trapezoidal fuzzy numbers and its application. *Applied Soft Computing, 28*, 276-284. Extraído desde 16/06/2016, de http://www.sciencedirect.com/science/article/pii/S1568494614006061.

SCHMUCKE, K. J. (1984). *Fuzzy Sets: Natural Language Computations, and Risk Analysis*. Computer Science Press, Incorporated.

VICENTE, E., MATEOS, A. Y JIMÉNEZ, A. (2013). *A new similarity function for generalized trapezoidal fuzzy numbers*. Paper presented at the International Conference on Artificial Intelligence and Soft Computing, Extraído desde 10/10/2016, de http://oa.upm.es/26121/1/26121mateos_INVE_MEM.pdf.

VICENTE, E. C., MATEOS, A. C. Y JIMÉNEZ , A. M. (2013). *Un enfoque borroso para el análisis y la gestión de riesgos en sistemas de información.* Tesis de maestría, Universidad Politécnica de Madrid. Extraído desde 25/9/216, de http://oa.upm.es/19054/2/TESIS_MASTER_ELOY_VICENTE_CESTERO.pdf.

WANG, Y.-M. Y ELHAG, T. M. (2006). Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment. *Expert Systems with Applications, 31*, 309-319.

XU, Z., SHANG, S., QIAN, W. Y SHU, W. (2010). A method for fuzzy risk analysis based on the new similarity of trapezoidal fuzzy numbers. *Expert Systems with Applications, 37*, 1920-1927.