Análisis de Riesgos en los Sistemas de Información.

Un Enfoque Difuso.

Risk Analysis in Information Systems.

A Fuzzy Approach.

E. Vicente, A. Mateos y A. Jiménez
Grupo de Análisis de Decisiones y Estadística.
Universidad Politécnica de Madrid. España.
e.vicentecestero@upm.es, {amateos,ajimenez}@fi.upm.es

Resumen— **En los modelos promovidos por las normativas internacionales de análisis de riesgos en los sistemas de información, los activos están interrelacionados entre sí, de modo que un ataque sobre uno de ellos se puede transmitir a lo largo de toda la red, llegando a alcanzar a los activos más valiosos para la organización. Es necesario entonces asignar el valor de todos los activos, así como las relaciones de dependencia directas e indirectas entre estos, o la probabilidad de materialización de una amenaza y la degradación que ésta puede provocar sobre los activos. Sin embargo, los expertos encargados de asignar tales valores, a menudo aportan información vaga e incierta, de modo que las técnicas difusas pueden ser muy útiles en este ámbito. Pero estas técnicas no están libres de ciertas dificultades, como la necesidad de uso de una aritmética adecuada al modelo o el establecimiento de medidas de similitud apropiadas. En este documento proponemos un tratamiento difuso para los modelos de análisis de riesgos promovidos por las metodologías internacionales, mediante el establecimiento de tales elementos.**

Abstract— **Assets are interrelated in risk analysis methodologies for information systems promoted by international standards. This means that an attack on one asset can be propagated through the network and threaten an organization's most valuable assets. It is necessary to valuate all assets, the direct and indirect asset dependencies, as well as the probability of threats and the resulting asset degradation. However, the experts in charge to assign such values often provide only vague and uncertain information. Fuzzy logic can be very helpful in such situation, but it is not free of some difficulties, such as the need of a proper arithmetic to the model under consideration or the establishment of appropriate similarity measures. Throughout this paper we propose a fuzzy treatment for risk analysis models promoted by international methodologies through the establishment of such elements.**

Keywords-component; análisis de riesgos, sistemas de información, números difusos trapezoidales)

I. INTRODUCCIÓN

Las normas promovidas por la Organización Internacional para la Estandarización (ISO) [10], [11] sobre seguridad de los sistemas de información (SI) sugieren metodologías de análisis y gestión de riesgos que contemplan tres fases fundamentales (Fig. 1).

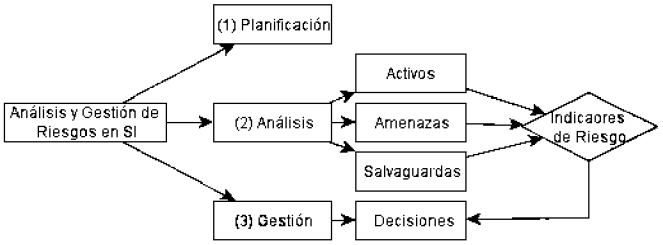

Figura 1: Proceso de Análisis y Gestión de Riesgos en los Sistemas de Información.

En la fase de análisis se determinan los activos que intervienen en el sistema de información, las relaciones entre éstos (dependencias), las amenazas a que están expuestos y, por último, las salvaguardas que se pueden implementar para hacer frente a esas amenazas. Los activos son los recursos del SI, o relacionados con éste, necesarios para que la Organización funcione correctamente y alcance los objetivos propuestos por su dirección. Pueden ser datos, aplicaciones software, instalaciones, hardware, servicios,...

Las dependencias entre los activos se suelen dar en términos de porcentaje, indicando con qué probabilidad los fallos de un activo pueden afectar a otro.

Generalmente, el valor total de los activos de una Organización se concentra en unos pocos elementos (activos terminales) que suelen ser de tipo "datos" o "servicios". El valor de estos activos se transmite al resto de activos a través de las relaciones de dependencia establecidas, de modo que los activos que no son terminales no tienen valor propio, sino que lo acumulan de los activos terminales. Sin embargo, las metodologías propuestas por las normas internacionales obvian la dificultad de asignar acertadamente las dependencias entre activos, así como el valor de los activos terminales o el impacto que provocaría sobre todo el sistema la materialización de una amenaza sobre un activo. Estas metodologías tampoco consideran la incertidumbre sobre estas valoraciones.

En este trabajo proponemos el uso de la Lógica Difusa en el

análisis de riesgos en los sistemas de información como solución a tales deficiencias. Para ello utilizamos una aritmética adecuada que extiende a los números difusos la metodología de valoración de dependencias entre los activos propuesta por las normas internacionales. Admitiendo además la inclusión de términos lingüísticos difusos en las componentes de valor de los activos terminales de la estructura básica del análisis de riesgos en los SI.

II. EVALUACIÓN DIFUSA DE LAS DEPENDENCIAS.

Consideremos el conjunto TF[0,1] de los números difusos trapezoidales con soporte en [0,1], es decir, dados por la tupla (a,b,c,d) con $0 \leq a \leq b \leq c \leq d \leq 1$, junto con una función que indica el grado de pertenencia a dicho número:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & si\ x < a \\ \dfrac{x-a}{b-a} & si\ a \leq x \leq b \\ 1 & si\ b \leq x \leq c \\ \dfrac{x-d}{c-d} & si\ c \leq x \leq d \\ 0 & si\ x > d \end{cases}$$

Consideremos en TF[0,1] la siguiente aritmética [25]:
Si $\tilde{A}_1 = (a_1, b_1, c_1, d_1)$ y $\tilde{A}_2 = (a_2, b_2, c_2, d_2)$ entonces:

$$\tilde{A}_1 \oplus \tilde{A}_2 = (a_1 + a_2 - a_1 a_2, b_1 + b_2 - b_1 b_2, c_1 + c_2 - $$

$$-c_1 c_2, d_1 + d_2 - d_1 d_2)$$

$$\tilde{A}_1 \otimes \tilde{A}_2 = (a_1 a_2, b_1 b_2, c_1 c_2, d_1 d_2)$$

$\oplus$ y $\otimes$ son dos leyes de composición interna en el conjunto TF[0,1] que verifican las propiedades conmutativa, asociativa y de existencia de elemento neutro .

Como indicábamos anteriormente, en los SI los activos están conectados mediante relaciones de dependencia, de modo que un fallo en un activo puede afectar al resto de los activos. Estas relaciones de dependencia llevan a una estructura como la que se puede ver en la Fig. 2, en la que el valor total del sistema se concentra en los activos terminales (información y datos, y servicios).

Diremos que el activo A_j depende del activo A_i , y lo denotaremos gráficamente por $A_i \rightarrow A_j$ si un fallo en A_i provoca un fallo en A_j con probabilidad $gr(A_i, A_j)$ a la que denominamos grado de dependencia de A_i sobre A_j.

Las metodologías oficiales sobre análisis de riesgo en los SI [7], [13], [16] asignan un porcentaje para indicar el grado de dependencia entre dos activos y, en ocasiones, proponen el uso de un valor booleano para indicar si existe o no tal dependencia, sin importar el grado en que ésta se produce. Frente a estas metodologías, nosotros proponemos el uso de números difusos trapezoidales, de modo que $gr(\widetilde{A_i, A_j}) \in TF[0,1]$. De este modo, los expertos pueden construir una escala de términos lingüísticos para asignar de forma intuitiva, y admitiendo imprecisión, la dependencia entre dos activos.

De la estructura general de dependencias del SI se sigue que la

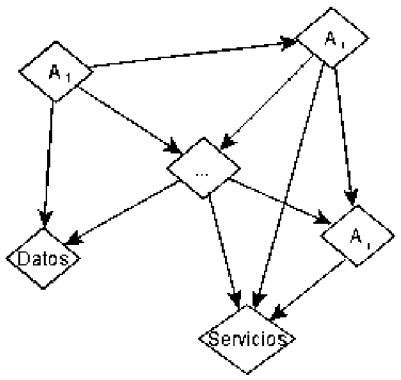

Figura 2: Estructura general de dependencias entre activos de un Sistema de Información.

dependencia entre activos no tiene por qué ser directa, sino que puede ser transitiva. Es decir, la transmisión de un fallo entre dos activos puede pasar de por activos intermedios, de modo que estamos interesados en calcular el grado de dependencia indirecto entre dos activos no consecutivos de la estructura general de dependencias. Denotaremos el grado de dependencia de A_i sobre A_k , a través de los activos intermedios A_j, como $Gr(\widetilde{A_i, A_k})$, y se calcula siguiendo el siguiente algoritmo (puede verse un ejemplo en la Sección IV):

Denotemos por $\mathbf{P} = \{P_1, P_2, \ldots, P_s\}$ el conjunto de caminos que conectan A_i con A_k.

A) Si todos los activos, salvo A_i y A_k, en los caminos de **P** están influidos por un solo activo entonces:

$$Gr(\widetilde{A_i, A_k}) = \oplus_{j=1}^{s}\ Gr(\widetilde{A_i, A_k}|P_j) \qquad (1)$$

donde

$Gr(\widetilde{A_i, A_k}|P_j) = gr(\widetilde{A_i, A_{j1}}) \otimes gr(\widetilde{A_{j1}, A_{j2}}) \otimes \ldots \otimes$

$\otimes gr(\widetilde{A_{jn}, A_k})$, con P_j: $(A_i \rightarrow A_{j1} \rightarrow \cdots \rightarrow A_{jn} \rightarrow A_k)$.

B) En otro caso, podemos asumir que los primeros r caminos de **P** están formados por caminos en los que cada activo está a su vez influido por un único activo, y los restantes $s - r$ caminos incluyen activos que están a su vez influidos por varios activos simultáneamente. Entonces para los r primeros caminos procedemos como en A) y denotamos por S el conjunto de los $s - r$ caminos restantes. En S procedemos de la siguiente manera:

   a. Consideramos el conjunto de activos no terminales de **S** influidos por dos o más activos. Denotamos por **I** a este conjunto. Sea **NI** el subconjunto de **I** de los activos que no están influidos por otro activo de **I** .
   b. Consideremos un activo A_r de **NI**. Entonces simplificamos el camino de **S** que incluye el activo A_r tomando : $A_i \rightarrow A_r \rightarrow \cdots \rightarrow A_k$ con $gr(\widetilde{A_i, A_r}) = Gr(\widetilde{A_i, A_r})$ que ha sido calculado en el paso anterior.
   c. Eliminamos los caminos repetido de **S** y man-

tenemos un único camino que agrupa a los anteriores.

d. Construimos de nuevo los conjuntos **I** y **NI** de **S**.
e. Si **NI** no es vacío volvemos al paso b. En otro caso, el algoritmo ha finalizado.
   Denotemos el conjunto de caminos resultante por $P = \{P'_1 \dots P'_m\}$ con $m \leq s - r$. Entonces el grado de dependencia de A_k, dado A_i es:

$$Gr(\widetilde{A_i, A_k}) = \oplus_{j=1}^{r} \; Gr(\widetilde{A_i, A_k}|P_j) \oplus \oplus_{l=1}^{m} \; Gr(\widetilde{A_i, A_k}|P'_l) \qquad (2)$$

El interés de las operaciones dadas radica en la acotación del conjunto de números difusos. Podemos asegurar que las operaciones entre términos lingüísticos difusos trapezoidales de una escala en [0,1] van a permanecer en TF[0,1], y mediante una función de similitud los resultados de estas operaciones se podrán traducir en uno de los términos lingüísticos de la escala. Además, la operación $\oplus$ es consistente con las metodologías establecidas de Análisis y Gestión de Riesgos como MAGERIT [13]–[15], propuesta por el Ministerio de Administraciones Públicas de España, y, junto con el algoritmo desarrollado anteriormente, permite interpretaciones en términos probabilísticos [20].

Dadas las operaciones $\oplus$ y $\otimes$, la metodología para derivar la dependencia de cada activo de los activos terminales consiste en los siguientes pasos:

- PASO 1: Se establece una escala de términos lingüísticos difusos.
- PASO 2: Se determina el grado de influencia de cada dos activos consecutivos en la estructura general de dependencias, estimando un término lingüísticos de la escala dada.
- PASO 3: Se determina el grado de influencia indirecto de los activos respecto de los activos terminales mediante las ecuaciones (1) y (2). Este grado de dependencia será un número difuso trapezoidal.

Un ejemplo de escala lingüísticos útil para este proceso se puede ver en la Tabla I.

III. VALORACIÓN DE LOS ACTIVOS.

La metodología MAGERIT define el valor de un activo como las pérdidas que tendríamos si prescindiéramos de dicho activo. Estas pérdidas pueden ser en concepto monetario, confianza de los usuarios, imagen de la organización...

Tabla I: Escala de términos lingüísticos y números difusos trapezoidales.

Término Lingüístico	Número Difuso
Muy Bajo (VL)	(0,0,0,0.05)
Bajo (L)	(0, 0.075, 0.125, 0.275)
Medio-Bajo (M-B)	(0.125, 0.275, 0.325, 0.475)
Medio (M)	(0.325, 0.475, 0.525, 0.675)
Medio-Alto (M-H)	(0.525, 0.675, 0.725, 0.875)
Alto (H)	(0.725, 0.875, 0.925, 1)
Muy Alto (V-H)	(0.925, 1, 1, 1)

Los activos tienen cinco componentes de valor [13]–[15]: *Confidencialidad* (¿Qué daño causaría que lo conociera quien no debe?), *integridad* (¿Qué perjuicio causaría que estuviera dañado o corrupto?), *autenticidad* (¿Qué perjuicio causaría no saber exactamente quién hace o ha hecho cada cosa?), *trazabilidad* (¿Qué daño causaría no saber a quién se le presta tal servicio?) y *disponibilidad* (¿Qué perjuicio causaría no tenerlo o no poder utilizarlo?).

Únicamente los activos terminales tienen valor propio. El resto de activos acumulan su valor a partir del valor de los activos terminales y las relaciones de dependencia.

Para considerar la imprecisión a la hora de valorar los activos terminales, al igual que para determinar las dependencias, vamos a recurrir al establecimiento de términos lingüísticos que representan números difusos.

Podemos escribir el valor propio en los activos terminales como $\tilde{v}_k = (\tilde{v}_{k1}, \tilde{v}_{k2}, \tilde{v}_{k3}, \tilde{v}_{k4}, \tilde{v}_{k5})$ donde $\tilde{v}_{ki}$ será un término lingüístico difuso asignado por un experto en la componente de valor (i)- ésima para el activo A_k.

El valor acumulado de un activo A_i respecto de los activos terminales A_k es:

$$\tilde{v}_{il} = \oplus_{k=1}^{n}((Gr(\widetilde{A_i, A_k})\tilde{v}_{kl} \qquad (3)$$

Para obtener el valor acumulado de los activos no terminales seguiremos, entonces, los siguientes pasos:

- PASO 4: Se estima el valor en cada componente de los activos terminales asignando un término lingüístico.
- PASO 5: Se calcula el valor acumulado en el resto de activos mediante la ecuación (3).

IV. LAS AMENAZAS.

Llamamos amenaza [13]–[15] a un evento que puede desencadenar un incidente en la Organización, produciendo daños materiales o pérdidas inmateriales en sus activos.

Tras haber valorado los activos, el siguiente paso en la metodología de análisis de riesgos es la valoración de las amenazas y la estimación de indicadores de impacto y riesgo sobre los activos. Para valorar las amenazas MAGERIT sugiere las medidas:

- Degradación: Perjuicio que la amenaza puede provocar sobre el activo.
- Frecuencia: Cada cuánto tiempo se materializa la amenaza.

La frecuencia se mide como el número medio de ocurrencias de la amenaza en un intervalo determinado de tiempo. Típicamente se estima sobre periodos anuales. En nuestro caso daremos términos lingüísticos difusos en lugar de porcentajes y probabilidades para indicar la degradación y la frecuencia. Una vez determinadas estas dos medidas se calculan los indicadores de impacto y riesgo que se definen a continuación.

Una amenaza es un vector $\vec{u} = (\tilde{f}, \tilde{d})$ cuyas componentes son la frecuencia y la degradación. Esta última, a su vez se puede dar en cada componente de valor.

Consideremos una amenaza sobre el activo A_j cuya degradación en cada componente viene dada por el vector

$\vec{\tilde{d}} = (\tilde{d}_1, \tilde{d}_2, \tilde{d}_3, \tilde{d}_4, \tilde{d}_5)$. Es decir, que la amenaza provoca una degradación de gravedad $\tilde{d}_i$ en la componente i- ésima del activo.

Cuando la amenaza se materializa, cada componente se verá afectada según la expresión

$$\tilde{I}_{ji} = \tilde{d}_i \otimes \tilde{v}_{ji} \quad (4)$$

$\tilde{I}$ es el impacto provocado sobre el activo atacado.

Para calcular el riesgo sobre este activo podemos utilizar la expresión

$$\tilde{R}_{ji} = \tilde{I}_{ji} \otimes f \quad (5)$$

Una vez calculado el impacto provocado por una amenaza materializada sobre un activo del sistema podemos calcular el impacto transmitido a los activos inferiores que dependen del activo atacado.

Si A_j es el activo sobre el que se ha materializado la amenaza y A_k un activo inferior cuyo grado de dependencia con A_j es $Gr(\widetilde{A_j, A_k})$ entonces la amenaza sobre el activo A_j provoca un impacto sobre A_k de $\tilde{I}_{ki} = Gr(\widetilde{A_j, A_k}) \otimes \tilde{d} \otimes \tilde{v}_{ki}$, de modo que el riesgo sobre el activo inferior será $\tilde{R}_{ki} = \tilde{I}_{ki} \otimes \tilde{f}$.

Por tanto, tras identificar las amenazas, su degradación y su frecuencia, los pasos que seguiremos para identificar el impacto y el riesgo sobre los activos atacados son:

- PASO 6: Se calculan los parámetros de impacto y riesgo en cada activo mediante las ecuaciones (4) y (5).
- PASO 7: Finalmente, el resultado difuso trapezoidal se asocia a uno de los términos lingüísticos de la escala dada mediante una función de similitud.

V. FUNCIÓN DE SIMILITUD.

Definimos el grado de similitud entre los números difusos trapezoidales $\tilde{A} = (a_1, a_2, a_3, a_4)$ y $\tilde{B} = (b_1, b_2, b_3, b_4)$ como:

- Si $\max\{(a_4 - a_1), (b_4 - b_1)\} \neq 0$ entonces:

$$S(\tilde{A}, \tilde{B}) = 1 - (1 - \alpha - \beta)\left(1 - \frac{\int_0^1 \mu_{\tilde{A}\cap\tilde{B}}(x)dx}{\int_0^1 \mu_{\tilde{A}\cup\tilde{B}}(x)dx}\right) - \alpha\frac{\sum|a_i - b_i|}{4} - \beta l_\infty[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})].$$

- En otro caso,

$$S(\tilde{A}, \tilde{B}) = 1 - \left(\frac{1-\alpha-\beta}{2} + \alpha\right)\frac{\sum|a_i - b_i|}{4} - \left(\frac{1-\alpha-\beta}{2} + \beta\right)|X_{\tilde{A}} - X_{\tilde{B}}|$$

donde

$$l_\infty[(x_1, y_1), (x_2, y_2)] = max\{|x_1 - x_2|, |y_1 - y_2|\}$$

$$\alpha + \beta < 1$$

$\mu_{\tilde{X}}(x)$ es la función de pertenencia de $\tilde{X}$.

$$\mu_{\tilde{A}\cap\tilde{B}}(x) = min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}$$

$$\mu_{\tilde{A}\cup\tilde{B}}(x) = max\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}$$

$$X_{\tilde{A}} = Y_{\tilde{A}}(a_3 + a_2) + (1 - Y_{\tilde{A}})(a_4 + a_1)$$

$$Y_{\tilde{A}} = \begin{cases} \dfrac{\left(\frac{a_3 - a_2}{a_4 - a_1}\right)}{6} & si\ a_4 - a_1 \neq 0 \\ 1/2 & si\ a_4 - a_1 = 0 \end{cases}$$

El punto $(X_{\tilde{A}}, Y_{\tilde{A}})$ es centro de gravedad del número difuso trapezoidal [4].

Proposición 1: $S(\tilde{A}, \tilde{B}) \in [0,1]\ \forall \tilde{A}, \tilde{B} \in TF[0,1]$

Demostración: Puesto que los pesos suman 1, basta ver que $\frac{\int_0^1 \mu_{\tilde{A}\cap\tilde{B}}(x)dx}{\mu_{\tilde{A}\cup\tilde{B}}(x)dx} \leq 1$, $\frac{\sum|a_i - b_i|}{4} \leq 1$, $l_\infty[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]$, lo cual es trivial.

Proposición 2: $S(\tilde{A}, \tilde{B}) = S(\tilde{B}, \tilde{A})\ \forall \tilde{A}, \tilde{B} \in TF[0,1]$

La demostración es trivial [21].

Proposición 3: $S(\tilde{A}, \tilde{B}) = 1 \Leftrightarrow \tilde{A} = \tilde{B}$.

Demostración: La implicación inversa es evidente. Veamos la implicación directa.

Si $\max\{(a_4 - a_1), (b_4 - b_1)\} \neq 0$, entonces

$$S(\tilde{A}, \tilde{B}) = 1 - (1 - \alpha - \beta)\left(1 - \frac{\int_0^1 \mu_{\tilde{A}\cap\tilde{B}}(x)dx}{\int_0^1 \mu_{\tilde{A}\cup\tilde{B}}(x)dx}\right) - \alpha\frac{\sum|a_i - b_i|}{4} - \beta l_\infty[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})] = 1 \Leftrightarrow$$

$$\Leftrightarrow (1 - \alpha - \beta)\left(1 - \frac{\int_0^1 \mu_{\tilde{A}\cap\tilde{B}}(x)dx}{\mu_{\tilde{A}\cup\tilde{B}}(x)dx}\right) - \alpha\frac{\sum|a_i - b_i|}{4} - \beta l_\infty[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})] = 0$$

Y como los tres sumandos son positivos o nulos, y $\alpha + \beta < 1$, deben ser:

$$\frac{\int_0^1 \mu_{\tilde{A}\cap\tilde{B}}(x)dx}{\int_0^1 \mu_{\tilde{A}\cup\tilde{B}}(x)dx} = 1, \frac{\sum|a_i - b_i|}{4} = l_\infty[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})] = 0$$

Por tanto $a_i = b_i\ \forall i = 1,2,3,4\ y\ \tilde{A} = \tilde{B}$.

Si $\max\{(a_4 - a_1), (b_4 - b_1)\} = 0$ la demostración es análoga.

Proposición 4: Si $\tilde{A} = (a, a, a, a)$ y $\tilde{B} = (b, b, b, b)$ entonces $S(\tilde{A}, \tilde{B}) = 1 - |a - b|$

Demostración: En estas hipótesis se tiene que $\max\{(a_4 - a_1), (b_4 - b_1)\} = 0$, por tanto:

$S(\tilde{A}, \tilde{B}) = 1 - \left(\frac{1-\alpha-\beta}{2} + \alpha\right)\frac{\sum|a_i - b_i|}{4} - \left(\frac{1-\alpha-\beta}{2} + \beta\right)|X_{\tilde{A}} - X_{\tilde{B}}| = 1 - \left(\frac{1-\alpha-\beta}{2} + \alpha\right)|a - b| - \left(\frac{1-\alpha-\beta}{2} + \beta\right)|a - b| = 1 - |a - b|$.

Estas propiedades garantizan la bondad de la función de similitud utilizada [21].

VI. EJEMPLO ILUSTRATIVO.

Consideremos la estructura de dependencias dada en la Fig. 3 en que el único activo terminal es A_6, y, por tanto, todo el valor del sistema se concentra en este activo. En esta figura se han señalado los grados de dependencia directos entre los activos, utilizando los términos lingüísticos de la Tabla I.

A. Dependencias indirectas sobre el activo terminal.

En primer lugar vamos a calcular el grado de influencia indirecto del activo A_1 sobre el activo A_6.

El conjunto de caminos que conectan A_1 con A_6 es:

$$\mathbf{P} = \{P_1: (A_1 \to A_2 \to A_6), P_2: (A_1 \to A_2 \to A_3 \to A_6)$$
$$P_3: (A_1 \to A_2 \to A_3 \to A_4 \to A_6), P_4: (A_1 \to A_3 \to A_6),$$
$$P_5: (A_1 \to A_3 \to A_4 \to A_6), P_6: (A_1 \to A_4 \to A_6),$$
$$P_7: (A_1 \to A_5 \to A_6)\}$$

El activo A_3 está influido por los activos A_1 y A_2, y A_4 está

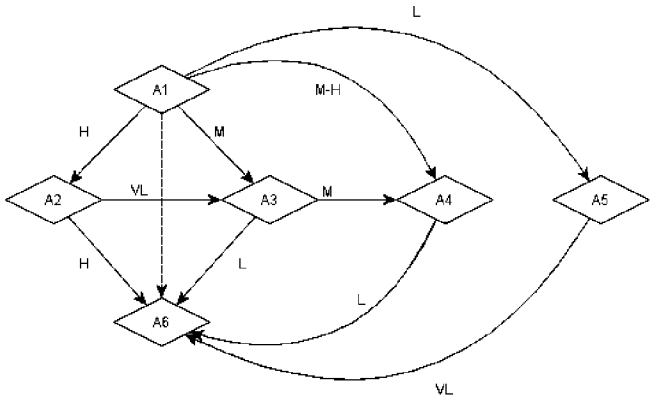

Figura 3: Estructura general de dependencias del ejemplo.

influido por A_1 y A_3. Por tanto, aplicamos el apartado B) del algoritmo de la Sección 2, con $r = 2$ y $\mathbf{S} = \{P_2, P_3, P_4, P_5, P_6\}$ y procedemos como sigue:

a. $I = \{A_3, A_4\}$ y $NI = \{A_3\}$
b. Seleccionamos A_3 entonces podemos simplificar P_2, P_3, P_4 y P_5 mediante $P'_2 = (A_1 \to A_3 \to A_6)$, $P'_3 = (A_1 \to A_3 \to A_4 \to A_6)$, $P'_4 = (A_1 \to A_3 \to A_6)$, $P'_5 = (A_1 \to A_3 \to A_4 \to A_6)$, respectivamente, con $gr(\widetilde{A_1, A_3}) = Gr(\widetilde{A_1, A_3}) = \big(gr(\widetilde{A_1, A_2}) \otimes gr(\widetilde{A_2, A_3})\big) \oplus gr(\widetilde{A_1, A_3})$.
c. $\mathbf{S} = \{P'_2, P'_3, P_6\}$ ya que $P'_2 = P'_4$ y $P'_3 = P'_5$.
d. $I = \{A_4\}$ y $NI = \{A_4\}$.
e. Ir al paso b.
b.2 Seleccionamos A_4 entonces simplificamos P'_3 y P_6 como $P''_3 = (A_1 \to A_4 \to A_6)$ y $P'_6 = (A_1 \to A_4 \to A_6)$, respectivamente, con $gr(\widetilde{A_1, A_4}) = Gr(\widetilde{A_1, A_4}) = \big(gr(\widetilde{A_1, A_3}) \otimes gr(\widetilde{A_3, A_4})\big) \oplus gr(\widetilde{A_1, A_4})$.
c.2 $\mathbf{S} = \{P'_2, P''_3\}$ ya que $P''_3 = P'_6$.
d.2 $I = \emptyset$ y $NI = \emptyset$.
e.2 El algoritmo finaliza ya que $NI = \emptyset$.

Finalmente $\mathbf{S} = \{P'_2, P''_3\}$ y el grado de dependencia de A_6 con respecto a A_1 es

$$Gr(\widetilde{A_1, A_6}) = Gr(\widetilde{A_1, A_6}|P_1) \oplus Gr(\widetilde{A_1, A_6}|P_7) \oplus Gr(\widetilde{A_1, A_6}|P'_2) \oplus Gr(\widetilde{A_1, A_6}|P''_3) = \big(gr(\widetilde{A_1, A_2}) \otimes gr(\widetilde{A_2, A_6})\big) \oplus \big(gr(\widetilde{A_1, A_5}) \otimes gr(\widetilde{A_5, A_6})\big) \oplus \big(gr(\widetilde{A_1, A_3}) \otimes gr(\widetilde{A_3, A_6})\big) \oplus \big(gr(\widetilde{A_1, A_4}) \oplus gr(\widetilde{A_1, A_6})\big).$$

Sustituyendo los arcos por los términos lingüísticos que indican los grados de dependencia obtenemos:

$$[[H \otimes [VL \otimes [(M \otimes L) \oplus L] \oplus H]] \oplus (L \otimes VL) \oplus \{[(H \otimes VL) \oplus M] \otimes \otimes [(M \otimes L) \oplus L]\} \oplus \{[[[(H \otimes VL) \oplus M] \otimes M] \oplus MH] \otimes L\} =$$

$$= [[(0.725, 0.875, 0.925, 1) \otimes [(0, 0, 0, 0.05) \otimes [((0.325, 0.475, 0.525, 0.675) \otimes (0, 0.075, 0.125, 0.275)) \oplus (0, 0.075, 0.125, 0.275)] \oplus (0.725, 0.875, 0.925, 1)]] \oplus ((0, 0.075, 0.125, 0.275) \otimes \{[(0.725, 0.875, 0.925, 1) \otimes (0, 0, 0, 0.05)) \otimes [((0.325, 0.475, 0.525, 0.675) \otimes (0, 0.075, 0.125, 0.275)) \oplus (0, 0.075, 0.125, 0.275)]] \oplus \{[[[((0.725, 0.875, 0.925, 1) \otimes (0, 0, 0, 0.05)) \oplus (0.325, 0.475, 0.525, 0.675)] \otimes (0.325, 0.475, 0.525, 0.675)] \oplus (0.525, 0.675, 0.725, 0.875)] \otimes (0, 0.075, 0.125, 0.275)\} =$$

$$= (0.525, 0.765, 0.855, 1) \oplus (0, 0, 0, 0.013) \oplus \oplus (0.325, 0.531, 0.611, 0.817 \oplus \oplus (0, 0.056, 0.1, 0.256)$$

$$= (0.679, 0.891, 0.949, 1).$$

Tabla II: Dependencias indirectas de cada activo con el activo terminal.

A_i	$Gr(\widetilde{A_i, A_6})$
A_1	(0.679, 0.891, 0.949, 1)
A_2	(0.725, 0.875, 0.925, 1)
A_3	(0, 0.107, 0.182, 0.409)
A_4	(0, 0.075, 0.125, 0.275)
A_5	(0, 0, 0, 0.05)

De forma análoga se calculan los grados de dependencia del resto de activos no terminales, dados en la Tabla II:

B. Valoración del activo A_1 a partir del valor del activoterminal y del grado de dependencia $Gr(\widetilde{A_1, A_6})$.

Supongamos que los expertos asignan un valor sobre A_6 dado por sus cinco componentes de $\tilde{v}_{6i} = (H, H, M, L, H)$. Entonces el valor acumulado sobre A_1 se calcula aplicando la ecuación (3), de donde obtenemos la Tabla III.

C. Amenazas. Indicadores de impacto y riesgo.

Consideremos una amenaza sobre el activo A_1 con una degradación $\vec{\tilde{d}} = (H, L, M, VL, M)$ y una frecuencia $\tilde{f} = M$. Entonces los indicadores de impacto y riesgo, que resultan de las ecuaciones (4) y (5) se pueden ver en las Tabla IV y V.

D. Términos lingüísticos para el riesgo.

Si ahora aplicamos la función de similitud sobre los pares formados por números difusos trapezoidales de la escala dada en la Tabla I y los indicadores de riesgo, podemos obtener un término lingüístico para expresar el riesgo en el activo atacado. Por ejemplo, la similitud del riesgo en la componente confidencialidad para cada término de la escala se puede ver en la Tabla VI, de donde se sigue que el riesgo sobre A_1 en dicha componente es Medio-Bajo.

VII. CONCLUSIONES.

Se ha desarrollado un modelo de análisis de riesgos en los sistemas de información basado en la metodología MAGERIT que incorpora conocimiento experto por medio de números difusos trapezoidales. El modelo difuso utiliza una aritmética adecuada basada en el cálculo de probabilidades para

establecer las dependencias entre los activos de información del sistema, de modo que el valor de dichos activos se determina a partir del valor de los activos terminales del sistema y de las dependencias con éstos. Al final del proceso de cálculo, por medio de una función de similitud de números difusos, se identifica el término adecuado de una escala lingüística previa para denotar el correspondiente índice de impacto y riesgo.

AGRADECIMIENTOS.

El desarrollo de este trabajo ha sido posible gracias a la financiación de la Comunidad Autónoma de Madrid a través del proyecto S-2009/ESP-1685 y del Ministerio de Ciencia y Tecnología del Gobierno de España a través del proyecto MYTM2011-28983-C03-03.

Tabla III: Valor acumulado de A_1 en cada componente.

Componente	$\tilde{v}_{1i}$
Confidencialidad	(0.492, 0.779, 0.877, 1)
Integridad	(0.492, 0.779, 0.877, 1)
Autenticidad	(0.22, 0.423, 0.498, 0.675)
Trazabilidad	(0, 0.066, 0.118, 0.275)
Disponibilidad	(0.492, 0.779, 0.877, 1)

Tabla IV: Indicadores de impacto sobre A_1.

Componente	Impacto
Confidencialidad	(0.35, 0.68, 0.81 1)
Integridad	(0, 0.05, 0.10, 0.27)
Autenticidad	(0.07, 0.2, 0.26, 0.45)
Trazabilidad	(0, 0, 0, 0.013)
Disponibilidad	(0.16, 0.37, 0.46, 0.67)

Tabla V: Indicadores de riesgo sobre A_1.

Componente	Riesgo
Confidencialidad	(0.11, 0.32, 0.43, 0.67)
Integridad	(0, 0.02, 0.05, 0.18)
Autenticidad	(0.02, 0.09, 0.13, 0.30)
Trazabilidad	(0, 0, 0, 0.01)
Disponibilidad	(0.05, 0.17, 0.24, 0.45)

Tabla VI: Similitud del riesgo calculado en la componente confidencialidad para el activo A_1.

Término lingüístico	Similitud
Muy Bajo (VL)	0.294
Bajo (L)	0.402
Medio-Bajo (M-B)	0.790
Medio (M)	0.588
Medio-Alto (M-H)	0.399
Alto (H)	0.176
Muy Alto (V-H)	0.076

REFERENCIAS.

[1] C. Alberts and A. Dorofee, "Managing Information Security Risks: The OCTAVE Approach". New York: Addison- Wesley, 2002.
[2] C. Alberts and A. Dorofee, "OCTAVE-s Method Implementation Guide Version 2.0". Pittsburgh: Canergie Mellon University, 2005.
[3] S.-M. Chen, "New methods for subjective mental workload assessment and fuzzy risk analysis", Cybernetics Syst., vol. 27, pp. 449–472, 1996.
[4] S.-J. Chen and S.-M. Chen, "A new method to measure the similarity between fuzzy numbers", Proc. 10th IEEE Int. Conf. Fuzzy Syst., pp. 208–214, 2001.
[5] ——, "Fuzzy risk analysis based on similarity measures of generalized fuzzy numbers", IEEE Trans. Fuzzy Syst, vol. 11, pp. 45–56, 2003.
[6] ——, "Fuzzy risk analysis based on the ranking of generalized trapezoidal fuzzy numbers", Appl Intell, vol. 26, pp. 1–11, 2007.
[7] CCTA Risk Analysis and Management Method (CRAMM), Version 5.0. London: Central Computing and Telecommunications Agency (CCTA), 2003.
[8] F. Herrera and L. Martínez., "A 2-tuple fuzzy linguistic representation model for computing with words", IEEE Trans. Fuzzy Syst., vol. 8, pp.746–752, 2000.
[9] C.H. Hsieh and S.H. Chen, "Similarity of generalized fuzzy numbers with graded mean integration representation", Proc. 8th Int. Fuzzy Syst. Assoc. World Congress, pp. 551–555, 1999.
[10] ISO/IEC 17799:2005, Information Technology - Security Techniques - Code of Practice for Information Security Management. Geneva: International Organization for Standardization, 2005.
[11] ISO/IEC 27005:2011, Information Technology - Security Techniques - Information Security risk Management. Geneva: International Organization for Standardization, 2005.
[12] H.S. Lee, "An optimal aggregation method for fuzzy opinions of group decision", Proc. 1999 IEEE Int. Conf. Syst., Man., and Cybernetics, pp. 314–319, 1999.
[13] F. López Crespo, M.A. Amutio-Gómez, J. Candau and J.A. Mañas, Methodology for Information Systems Risk. Analysis and Management (MAGERIT version 2). Book I-The Method. Madrid: Ministerio de Administraciones Públicas, 2006a.
[14] ——, Methodology for Information Systems Risk Analysis and Management (MAGERIT version 2). Book II-Catalogue of Elements. Madrid: Ministerio de Administraciones Públicas, 2006b.
[15] ——, Methodology for Information Systems Risk Analysis and Management (MAGERIT version 2). Book III-The Techniques. Madrid: Ministerio de Administraciones Públicas, 2006c.
[16] Mehari 2010 - Risk Analysis and Treatment Guide. Paris: Club de la Sécurité de l'Information Francais (CSIF), 2007.
[17] A. Méndez Barco and J.A. Mañas, Manual del Usuario Pilar Basic versión 5.1. Madrid: Centro Criptológico Nacional, 2011.
[18] G. Stoneburner and A. Gougen, NIST 800-30 Risk Management. Guide for Information Technology Systems. Gaithersburg: National Institute of Standard and Technology, 2002.
[19] A. Syalim, Y. Hori and K. Sakurai, "Comparison of risk analysis methods: Mehari, Magerit, NIST800-30 and Microsoft's Security Management Guide", Proc. Int. Conf. Availability, Reliability and Security, pp. 726–735, 2009.
[20] E. Vicente, A. Jiménez and A. Mateos. "A Fuzzy Approach to Risk Analysis in Information Systems". 2nd Internationa Conference on Operations Research and Enterprise Systems. Barcelona 2013.
[21] E. Vicente, A. Mateos and A. Jiménez, "A New Similarity Function for Generalized Trapezoidal Fuzzy Numbers", 12th International Conference on Artificial Intelligence and Soft Computing. Zakopane 2013. To appear.
[22] J. Wang and J. Hao, "A new version of 2-tuple fuzzy linguistic representation model for computing with words", IEEE Trans. Fuzzy Syst., vol. 14, pp. 435–445, 2006.
[23] S. H. Wei and S. M. Chen, "A new approach for fuzzy risk analysis based on similarity measures of generalized fuzzy number", Expert Syst. Appl., vol. 36, pp. 589–598, 2009.
[24] Z. S. Xu, "A method based on linguistic aggregation operators for group decision making with linguistic preference relations", Inform. Sci., vol. 166, pp.19–30, 2004.
[25] Z. Xu, S. Shang, W. Qian and W. Shu, "A method or fuzzy risk analysis based on the new similarity of trapezoidal fuzzy numbers", Expert Syst. Appl., vol. 37, pp. 1920–1927, 2010.
[26] L. A. Zadeh, "Fuzzy sets", Inform. Control, vol. 8, pp. 338–353, 1965.