Entre Ciencia e Ingeniería, ISSN 1909-8367
Año 9 No. 17 - Primer Semestre de 2015, página 71 - 80

Definición de un modelo de medición de análisis de riesgos de la seguridad de la información aplicando lógica difusa y sistemas basados en el conocimiento[1]

Definition of a model for measuring risk analysis of security information applying fuzzy logic and systems based on knowledge

A.A.Angarita, C.A.Tabares y J.I.Rios

Recibido Abril 10 de 2015 – Aceptado Mayo 29 de 2015

Resumen **- La información es el activo más valioso de una organización, y por lo tanto, es necesario implementar técnicas cada vez más sofisticadas para protegerla. Dentro de las diferentes técnicas que pueden utilizarse para medir el análisis de riesgos de seguridad de la información se encuentran la lógica difusa y los sistemas basados en el conocimiento. La aplicación de las propiedades de la lógica difusa permite realizar el análisis de riesgos de seguridad de la información a partir de los criterios y la experiencia de especialistas, creándose un sistema difuso que tiene en cuenta la subjetividad del análisis y a su vez constituye una herramienta de sencilla utilización.**

Palabras Clave **- Lógica difusa, impacto, sistemas basados en el conocimiento, Matlab, prototipo, riesgo, amenaza, vulnerabilidad.**

Abstract **- Information is the most valuable asset of an organization, and therefore, it is necessary to implement increasingly sophisticated techniques to protect it. Among the different techniques that can be used to measure risk analysis information, we can count security fuzzy logic and knowledge-based systems. The application of the properties of fuzzy logic allows risk analysis of information security from the approaches and experiences of specialists, creating a fuzzy system that takes into account the subjectivity of analysis and in turn it is a tool of simple use.**

Key Words **- Fuzzy logic, impact, knowledge-based systems, Matlab, prototype, risk, threat, vulnerability.**

I. Introducción

En un artículo publicado por Josh James[2] el 23 de abril de 2014 se revela un estudio titulado DATA NEVER SLEEPS 2.0, el cual tiene como objetivo principal mostrar la cantidad de datos que son generados cada minuto en internet. Dentro de los datos más sorprendentes se destacan los siguientes:

- Los usuarios de YouTube[3] suben 72 horas de nuevos videos.
- Se envían 204.000.000 correos electrónicos
- Google[4] recibe 4.000.000 de búsquedas
- Los usuarios de Facebook[5] comparten 2.460.000 piezas de contenido
- Se envían 277.000 tweets en Twitter[6].

[1]Producto derivado del proyecto de investigación "Definición de un modelo de análisis de riesgos de la seguridad de la información aplicando lógica difusa y sistemas basados en el conocimiento". Presentado por el grupo de investigación ADA de la Universidad Tecnológica de Pereira.

J.I.Rios. Director Maestría Ingeniería de Sistemas y Computación, de la Universidad Tecnológica de Pereira, Pereira (Colombia), email: jirios@utp.edu.co

A.A.Angarita. Administrador de plataforma tecnológica, de la empresa Aguas y Aguas de Pereira, Pereira (Colombia), email: angarco@gmail.com

C.A.Tabares. Catedrático auxiliar Facultad de Ingenierías, de la Universidad Tecnológica de Pereira, Pereira (Colombia), email: ceautabares@utp.edu.co

Es innegable que un minuto en informática es como hablar de un año-luz en astronomía, y que con toda la información que circula por la red de datos mundial es necesario definir estrategias para garantizar la seguridad de ésta.

Distintas organizaciones de gran reconocimiento internacional han creado modelos que definen una serie de pasos para que las empresas implementen sistemas de seguridad de la información basados en el análisis de riesgos. La implementación de estos sistemas implica la asignación de pesos específicos para las amenazas y las vulnerabilidades, siendo ésta una responsabilidad del experto.

Sin embargo, los expertos encargados de asignar tales valores a menudo aportan únicamente información imprecisa, de modo que las técnicas difusas pueden ser muy útiles en este ámbito.

Con el fin de dar un tratamiento computacional a esta información imprecisa, es necesario proveer a los expertos de un método con el que puedan expresar sus juicios o sus valoraciones sobre los activos de información en forma de números difusos, evitando sesgos informativos. Una vez obtenidos tales valores, se construyen los algoritmos que permiten establecer indicadores de impacto y riesgo para las amenazas que se ciernen sobre los activos de información, y finalmente se proponen conjuntos óptimos de salvaguardas y controles para reducir el nivel de riesgo aceptable por la organización.

En este trabajo se plantea un modelo que permite realizar este proceso partiendo de los conceptos básicos de lógica borrosa y de las metodologías internacionales de análisis y gestión de riesgos en los sistemas de información.

El documento consta de una definición del problema, la justificación del trabajo, definición de los conceptos esenciales, la metodología utilizada para el desarrollo del proyecto, la implementación del modelo con los resultados obtenidos en un caso de estudio y las conclusiones finales.

II. Definición del problema

En el proceso de análisis de riesgos, independientemente del ámbito donde se realice y la metodología que se utilice, se hace necesario definir y utilizar escalas cualitativas que dependen generalmente del enfoque subjetivo del experto a cargo (ej. bajo, medio, alto),en las que se puede incluir una variedad de información descriptiva, incluyendo datos y opiniones, con el fin de obtener la medición del nivel de riesgo, el impacto y la probabilidad de que una amenaza aproveche una vulnerabilidad.

Como resultado de lo anterior, los datos que se obtienen no brindan la suficiente precisión, y por consiguiente no es posible realizar la estimación del riesgo de una manera más efectiva y concisa.

La base para la priorización de riesgos es la Matriz de Riesgos, la cual se compone de un eje de la consecuencia o el impacto que genera la materialización de una amenaza y otro de la Frecuencia o probabilidad de que una amenaza se materialice. El producto del Impacto y la Probabilidad proporciona una medida de riesgo. Sin embargo, aunque este es el enfoque estándar utilizado en la mayoría de las organizaciones, pareciera no proporcionar resultados concluyentes y eficaces, que reflejen el esfuerzo invertido o satisfagan las necesidades de los procesos de decisión subsecuente.

Lo anteriormente expuesto conlleva a que en gran cantidad de ocasiones los expertos no posean la capacidad o cuenten con metodologías que les permitan definir valores concretos para las diferentes alternativas entre las cuales se debe decidir así como los pesos de los diferentes criterios de decisión o las probabilidades de ocurrencia de los distintos escenarios.

III. Justificación

En el proceso de análisis de riesgos, independientemente del ámbito donde se realice y la metodología que se utilice, se hace necesario definir y utilizar escalas cualitativas que dependen generalmente del enfoque subjetivo del experto a cargo (ej. bajo, medio, alto), en las que se puede incluir una variedad de información descriptiva, incluyendo datos y opiniones, con el fin de obtener la medición del nivel de riesgo, el impacto y la probabilidad de que una amenaza aproveche una vulnerabilidad. Como resultado de lo anterior, los datos que se obtienen no brindan la suficiente precisión, y por consiguiente no es posible realizar la estimación del riesgo de una manera más efectiva y concisa.

Es por eso que se plantea definir un modelo de medición de análisis de riesgos de la seguridad de la información aplicando lógica difusa y sistemas basados en el conocimiento, que permita entregar a los expertos en el tema, una herramienta que les facilite expresar sus juicios probabilísticos o sus valores sobre los activos de información en forma más exacta evitando así la ambigüedad y la declaración de conceptos subjetivos y permitiendo el tratamiento computacional de los mismos.

IV. Seguridad de La Información

La seguridad de la información es la preservación de los principios básicos de la confidencialidad, integridad y

[2] Founder, CEO & Chairman of DOMO

[3]Sitio web en el cual los usuarios pueden subir y compartir videos

[4]Empresa multinacional estadounidense especializada en productos y servicios relacionados con internet

[5]Sitio web de redes sociales creado en el año 2004 por Mark Zuckerberg

[6]Empresa estadounidense especializada en servicios de publicación instantánea de mensajes

disponibilidad de la misma y de los sistemas implicados en su tratamiento. Estos tres principios básicos se definen de la siguiente manera:

Confidencialidad: Acceso a la información por parte únicamente de quienes estén autorizados.

Integridad: Mantenimiento de la exactitud y completitud de la información y sus métodos de proceso.

Disponibilidad: Acceso a la información y los sistemas de tratamiento de la misma por parte de los usuarios autorizados cuando lo requieran.

En la seguridad de la información, no sólo intervienen los aspectos tecnológicos, sino también los procesos, los ambientes (centro de cómputo, ubicación de oficinas) y principalmente las personas.

V. Lógica Difusay Análisis de Riesgos

En los modelos promovidos por las normativas internacionales de análisis y gestión de riesgos en sistemas de información, los activos están relacionados 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 éstos, 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 únicamente información imprecisa, de modo que las técnicas borrosas pueden ser útiles en este ámbito. [7]

Para poder dar un tratamiento computacional a esta información imprecisa, es necesario proveer a los expertos de un método con el que puedan expresar sus juicios probabilísticos o sus valoraciones sobre los activos de información, en forma de números difusos y evitando sesgos informativos. Una vez obtenidos tales valores, se construyen algoritmos que nos permiten establecer indicadores de impacto y riesgo para las amenazas que se ciernen sobre los activos de información, y finalmente se proponer conjuntos óptimos de salvaguardas y controles para reducir el riesgo a un nivel asumible.

A pesar del hecho de demostrar este gran potencial, la lógica difusa no ha encontrado un lugar aún en el proceso de priorización de riesgos, probablemente debido a la negativa de apartarse de los métodos ya existentes y los requerimientos conceptuales de la formulación del problema.

Teniendo esto en mente, los conceptos propuestos por Adam Markowsky y Sam Mannan [9] de una matriz de riesgo difusa (FRM) fueron seleccionados como la etapa inicial para el desarrollo del índice de priorización de riesgo. Este enfoque resulta extremadamente práctico, al tener como punto de partida la matriz tradicional de riesgo (TRM) hereda todas sus capacidades ampliamente revisadas en la literatura; mientras se integra fácilmente con los esquemas actuales de clasificación y priorización de riesgos. [7]

VI. Metodología

La metodología a seguir se muestra en la Fig.1.

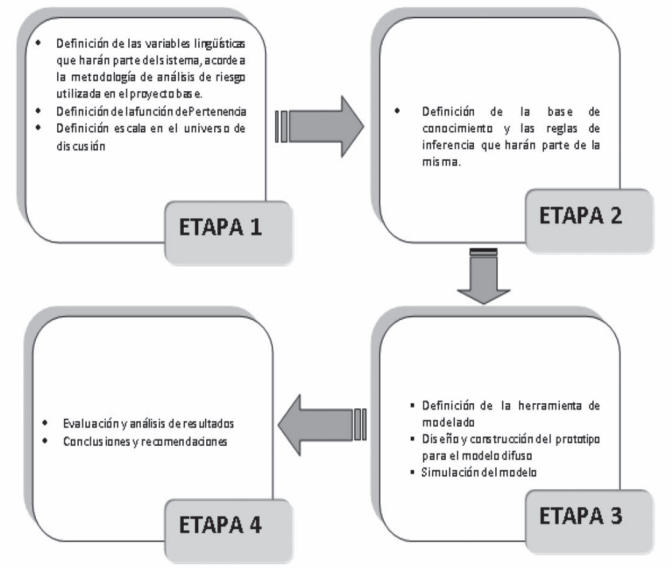

Fig.1. Metodología del Proyecto

Implementación Del Modelo

Las entradas del sistema son los valores asumidos para las diferentes probabilidades de que las amenazas se puedan materializar para un activo de información en particular; así mismo, el impacto que se puede generar en el momento que se materialice alguna de las amenazas. Éstas son tomadas del documento [3].

Teniendo en cuenta dicho documento, las variables lingüísticas de entrada definidas para el prototipo son:

Daño Físico:	DF
Compromiso de la Información:	CI
Fallas Técnicas:	FT
Acciones no Autorizadas:	ANA
Impacto:	I

Para valorar la Probabilidad de que una amenaza se materialice, logrando de esta manera afectar la disponibilidad, confidencialidad, integridad y/o autenticidad de la información, se toma en cuenta la descripción contenida en la Tabla I.

Tabla I
Valoración de La Probabilidad

	VALOR	DESCRIPCIÓN
1	Insignificante (I)	No existen condiciones que impliquen que el hecho se presente.
2	Baja (B)	Existen condiciones que hacen muy lejana la posibilidad de que el hecho se presente.
3	Mediana (M)	Existen condiciones que hacen poco probable un hecho en el corto plazo pero que no son suficientes para evitarlo en el largo plazo.
4	Alta (A)	La realización del hecho es inminente. No existen condiciones internas y externas que impidan el desarrollo del hecho.

Para valorar el impacto se toma como base la información contenida en la Tabla II.

Tabla II
Valoración del Impacto

	VALOR	DESCRIPCIÓN
1	Insignificante (I)	No causa ningún tipo de impacto o daño al Área o la organización.
2	Bajo (B)	Causa daño aislado, que no perjudica a ningún componente del Área o de la organización.
3	Mediano (M)	Provoca la desarticulación de un componente del Área o de la de la organización. Si no se atiende a tiempo, a largo plazo puede provocar la desarticulación de la organización.
4	Alto (A)	En el corto plazo desmoviliza o desarticula a la organización.

La variable lingüística de salida de este prototipo es denominada Nivel de Riesgo (NR), la cual es la representación del nivel de riesgo de seguridad del activo de información evaluado con las etiquetas lingüísticas, de acuerdo con la información contenida en la Tabla III.

Tabla III
Nivel de Riesgo

VALOR	DESCRIPCIÓN
Alto (A)	Se requiere de acciones inmediatas.
Medio (M)	Se requiere de acciones a mediano plazo.
Bajo (B)	Se requiere de acciones a largo plazo

Para el proceso de validación se tomarán datos reales emitidos por expertos en el tema, los cuales se introducirán en el campo correspondiente de la interfaz hecha para el prototipo de evaluación de riesgos de seguridad de la información, todos los datos se supondrán acorde a la realidad expresada mediante el concepto del experto, arrojando así un resultado numérico que indicará la calificación del nivel de riesgo en el que se encuentra el activo evaluado.

Una vez definidas las variables lingüísticas de entrada y de salida, se analizan las diferentes funciones de pertenencia existentes y se elige la Función de pertenencia TRIANGULAR.

Posteriormente se definen los conjuntos difusos de las variables de entrada que en este caso son los valores que se tienen para la probabilidad de que una amenaza se materialice, así como el valor del impacto que se tiene si dicha amenaza se llegara a materializar (Insignificante, Bajo, Medio y Alto).

Así, el universo del discurso es igual para todas las variables de entrada que se han definido para el prototipo, por lo tanto, todas las variables de entrada tienen las mismas etiquetas lingüísticas y funciones de pertenencia. La Fig. 2 ilustra las funciones de pertenencia de las etiquetas lingüísticas de las variables de entrada del modelo.

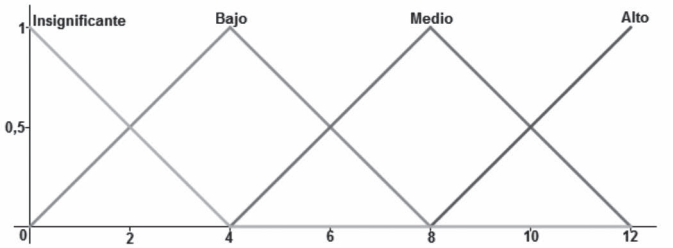

Fig.2. Función de pertenencia de las etiquetas lingüísticas de las variables de entrada del modelo.

Por último se definen los conjuntos difusos para la variable de salida del prototipo en este caso el Nivel de Riesgo (NR). El universo de discurso de esta variable está comprendido entre 0 y 16, que corresponden a los valores mínimo y máximo que puede tomar el nivel de riesgo acorde a la tabla definida en el documento [3], y que se describen en la Fig. 3.

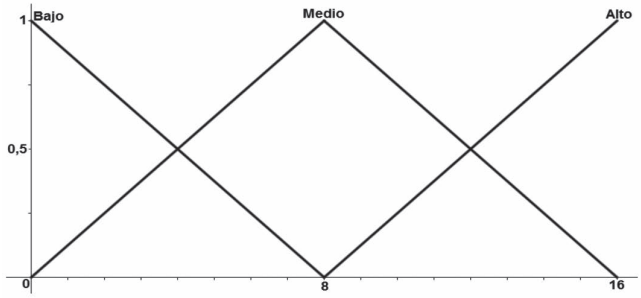

Fig.3. Función de pertenencia de las etiquetas lingüísticas de la variable de salida del modelo.

- Marco de representación de la solución

En el prototipo se tendrán en cuenta las diferentes variables de entrada para proporcionar el conocimiento necesario al prototipo y obtener una respuesta deseada acorde al fin propuesto, para tener una mejor idea de las variables que se utilizan en éste, se representarán La Fig. 4.

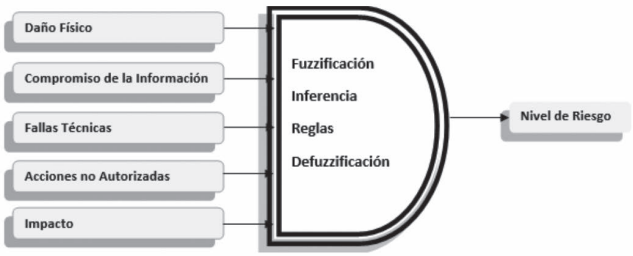

Fig.4. Representación de las variables lingüísticas y los procesos que se realizan en el prototipo.

Las variables de entrada provienen del número de amenazas que se pueden presentar sobre un activo de información, así como el impacto que se genera si dichas amenazas se materializan. Estos valores son los que entrarán en el proceso del motor de lógica difusa, para así arrojar un resultado numérico el cual posicionará el nivel de riesgo de seguridad en el que se encuentra el activo de información en un rango contemplado y le dará la calificación de alto, medio o bajo, según sea considerado.

- Para dar la calificación al Nivel de Riesgo se propone un rango de valores, los cuales son:

- Si el valor numérico está entre 0 y 8 el nivel de riesgo tendrá un valor entre bajo y medio, según el área que ocupe en las funciones de pertenencia.

- Si el valor numérico está entre 0 y 16 la calificación tendrá un valor entre bajo, medio y alto, según el área que ocupe en las funciones de pertenencia.

- Si el valor numérico está entre 8 y 16 la calificación tendrá un valor entre medio y alto, según el área que ocupe en las funciones de pertenencia.

Las reglas fueron construidas mediante conceptos de expertos en el tema, teniendo en cuenta la implicación borrosa por la regla del mínimo o la implicación de Mamdani[7], además de utiliza las operaciones entre conjuntos como la intersección con sus diferentes fórmulas, con el fin de analizar adecuadamente las entradas al sistema.

Para el proceso de inferencia se utilizan las diferentes reglas definidas en el motor difuso, que provienen de la combinación de los conjuntos difusos de las diferentes variables de entrada, las cuales arrojan un resultado difuso (variable de salida).

Para definir esta salida llamada NIVEL DE RIESGO, se asume que todas las variables de las reglas tienen el mismo "peso", sin dar prioridad a ninguna para evitar confusión y conflictos en la toma de decisiones, dado que no se cuenta con estudios estadísticos y en el alcance del proyecto no se contempla realizar este estudio. Por lo tanto se suponen los rangos y los datos para evitar entrar a ambigüedades y solo se centra en demostrar que el uso de la lógica difusa es de gran ayuda.

• Salida del prototipo (decodificación)

La salida del prototipo se realiza a través del procedimiento del método del centroide, donde a partir de un valor difuso que sale del proceso de inferencia se obtiene un valor numérico, el cual es visto por el usuario final.

• Definición y cálculo de centroides

En matemáticas, los centroides de una figura bidimensional se refieren al punto en el cual todas las líneas de la figura correspondiente se intersectan unas con otras, de tal manera que dividen la figura en dos partes iguales en los momentos equivalentes. Si se establece físicamente, un centroide se refiere al centro del objeto geométrico.

En el capítulo de análisis de resultados se requiere conocer el centroide de un triángulo rectángulo y de un rectángulo.

El centroide de un triángulo rectángulo está ubicado a un tercio de su base y a un tercio de su altura.

El centroide de un rectángulo está ubicado a un medio de su base y a un medio de su altura.

Son muchos los lenguajes que proporcionan herramientas que permiten trabajar con lógica difusa, teniendo todos ellos características diferenciadoras, las cuales permiten tener opciones para escoger cual es la herramienta más adecuada para el desarrollo planteado en el documento, para esto se tienen en cuenta ventajas y desventajas de las diferentes opciones. Luego de una exhaustiva y cuidadosa investigación sobre las herramientas posibles a utilizar, la opción seleccionada es FuzzyLogic de MATLAB.

A. Diseño y construcción del prototipo para el modelo difuso

Definición de restricciones y comportamiento del sistema.

Restricciones:

- El prototipo no mostrará recomendaciones.
- El prototipo no analizará los diferentes componentes de confidencialidad, integridad y disponibilidad de la información en el análisis de riesgos.
- El prototipo no tomará decisiones finales sobre la mitigación de riesgos de seguridad de la información, sólo mostrará la el nivel de riesgo para el activo de información definido en el modelo.
- Los datos cargados al prototipo serán obtenidos de las siguientes fuentes:
  1. Probabilidad de que una amenaza se materialice.
  2. Impacto al que hubiere lugar si una amenaza se materializa.
- Los datos se cargarán al prototipo de manera manual.
- Solamente se puede hacer el ingreso de los datos y su respectiva evaluación de riesgo a un activo de información a la vez.
- El prototipo no mostrará valores probabilísticos de las evaluaciones de riesgos en la seguridad de la información realizadas, sino valores de pertenencia a la variable de salida.

[7]En honor a Ebrahim Mamdani (Junio 1942, Enero 2010), profesor emérito del Departamento de Ingeniería Eléctrica y Electrónica del Imperial College of London. Desarrolló el sistema difuso que lleva su nombre.

El prototipo funciona basado en la calificación de la probabilidad de que una amenaza pueda ocurrir y el valor del impacto que esta amenaza generaría en caso de materializarse. Estos valores de entrada son convertidos en valores difusos mediante las funciones de pertenencia triangulares, luego éstos pasan al motor de inferencia donde gracias a las reglas definidas, se obtiene un resultado difuso, el cual pasará al proceso de desfuzzyficación, obteniendo así un valor numérico que permitirá obtener el nivel de riesgo de seguridad de la información para el activo evaluado. Estos valores de salida también tienen una función de pertenencia la cual es triangular.

B. Simulación del Modelo

• Explicación de la interfaz del prototipo

Dado que el presente proyecto no busca la construcción de un sistema completo para el proceso de evaluación de riesgos de seguridad de la Información y sólo apunta a validar la hipótesis de que la lógica difusa es una excelente herramienta para apoyar dicho proceso; se desarrolló un prototipo con una vista simple y sencilla que permite validar de manera clara el poder de la lógica difusa y sobre todo la facilidad en el manejo de un controlador difuso una vez se tiene construido.

La construcción de la vista fue basada en la tipificación de amenazas que se pueden encontrar en los activos de información, así como el nivel de impacto que se puede tener dado el caso de que estas amenazas se materialicen.

De esta manera, la vista consta de una columna de espacios editables de texto (Amenazas e impacto) en donde se ingresan de manera manual los valores acorde al concepto del experto.

Una vez ingresados los valores de las amenazas y el impacto, se debe dar clic en el botón "Evaluar" el cual cumple la labor de enviarle al controlador difuso los valores de la columna (Amenazas e Impacto) para su análisis y posteriormente muestra en pantalla, en un cuadro de texto, el resultado del análisis de los valores, es decir, el NIVEL DE RIESGO.

La interface posee algunos flujos de control que permiten la conservación de los siguientes requerimientos funcionales del prototipo:

Requerimiento 1. El valor de los campos para las Probabilidades de Amenaza e Impacto deben estar entre un rango de 0 a 12.

Control para el requerimiento 1: La interfaz cuenta con avisos de ERROR en la parte derecha, los cuales son mostrados al usuario en el momento de ingresar valores que no están en el rango especificado.

Requerimiento 2. En los campos donde no se digite ningún valor, se asumirá un valor mínimo de CERO.

Control para el requerimiento 2: Se define la función "CajonVacio", la cual valida si la caja de texto no tiene ningún valor, y en su defecto coloca el valor 0 cero para esta caja de texto:

```
function cajonVacio(hObject)
if (strcmp(get(hObject,'String'),{''}))
set(hObject, 'String', '0');
end
```

Si no existe ningún inconveniente y los valores para la probabilidad de amenaza y los valores del impacto están en los rangos establecidos, se utiliza el botón "Evaluar Riesgo" y se procede con la obtención del nivel de riesgo para el activo evaluado.

Por último, el botón llamado "Limpiar" permite reiniciar los valores de todas las cajas de texto y del cuadro de texto "Nivel de Riesgo" para volver a realizar el proceso de evaluación a un activo especifico.

C. Ejecución de Pruebas y Validación del Prototipo

Las pruebas para validar el prototipo se realizaron con datos reales entregados por expertos en el tema de seguridad de la información y en algunos casos con datos irreales, ya que lo que se pretende es dar a conocer el funcionamiento de un prototipo basado en lógica difusa que puede ser aplicado a la evaluación de riesgos de seguridad de la información.

Para poder determinar la validez de la hipótesis planteada: "Aplicando lógica difusa se puede generar un modelo para la evaluación del análisis de riesgos de seguridad de la información", es necesario realizar algunas pruebas para observar el comportamiento del prototipo construido y poder apreciar los resultados que éste arroja.

Para un mejor entendimiento de las pruebas, remitirse al apartado No 8 [3] donde se explica el proceso de evaluación de riesgo, en el que básicamente se tienen como fuentes los activos de información, las amenazas que se pueden materializar sobre estos activos y el impacto que genera en la Empresa si llegase a ocurrir esta materialización. Estos tres elementos permitirán hacer una evaluación de riesgo de seguridad de la información.

Una vez se ingresan los valores de las variables antes mencionadas, se puede iniciar la evaluación del riesgo usando el prototipo diseñado, el cual analiza los datos e información ingresada, compara con las reglas de inferencia internas del controlador y arroja el nivel de riesgo para el activo de información evaluado.

Teniendo claro lo anterior y los activos de información definidos en el documento [3], se tomará la matriz de riesgos definida, y se escogerán algunos activos de información para su análisis y evaluación en el prototipo.

La matriz de riesgos se incluye en el Anexo A, la cual hace parte del documento [3], el cual sirvió como base para tipificar todas las amenazas que pueden materializarse en un sistema de información.

NOTA: Cabe recordar que las amenazas fueron tipificadas y simplificadas para efectos de poder realizar el prototipo inicial.

Se da inicio a las pruebas del prototipo construido:

PRUEBA UNO: Se toma el activo de Información "Servidores"

Se efectúa la calificación de expertos acorde a los datos que se deben ingresar al prototipo:

La valoración de amenazas para el activo de Información Servidores se describe en la Tabla IV.

Tabla IV
Valoración de Amenazas para El Activo Servidores

AMENAZA	VALORACIÓN
Daño Físico	3
Compromiso de la Información	3
Fallas Técnicas	3
Acciones no autorizadas	4

La valoración del Impacto para el activo de Información Servidores es igual a 4.

La Fig. 5 muestra el resultado obtenido al ingresar los datos anteriores al prototipo.

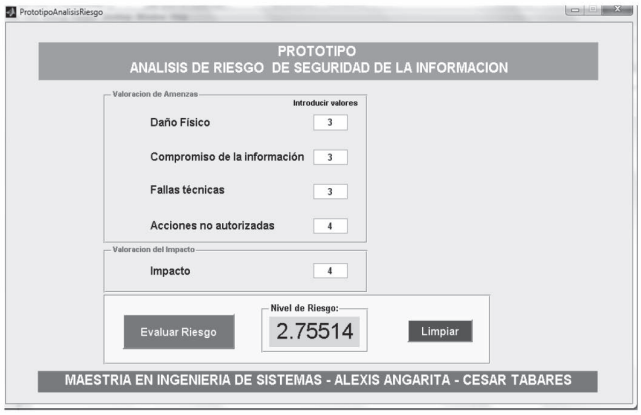

Fig.5. Resultado del Prototipo con el vector de entrada [3, 3, 3, 4, 4]

A continuación se analizarán las reglas que son estimuladas en la fase de inferencia dentro del controlador difuso, según las entradas expuestas anteriormente.

Acorde a la prueba anterior se tiene el vector de entrada [3, 3, 3, 4, 4] en el controlador, con el cual se estimulan las reglas mostradas en la Tabla V.

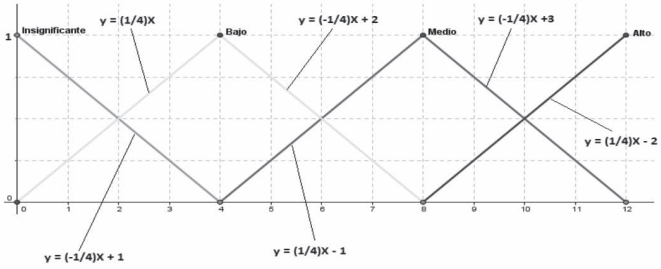

Fig.6. Ecuaciones de las funciones de pertenencia de las etiquetas lingüísticas de las variables de entrada

NOTA: Un valor de 4 para la etiqueta lingüística de las diferentes variables de entrada del prototipo, pertenece en mayor o menor grado a las funciones de pertenencia INSIGNIFICANTE, BAJO y MEDIO. Mientras que un valor de 3 para la misma etiqueta lingüística sólo pertenece a las funciones de pertenencia INSIGNIFICANTE y BAJO.

A partir de las ecuaciones de la Fig.6, se obtienen los valores del grado de pertenencia para cada una de las entradas del prototipo.

Tabla V
Reglas que satisfacen vector de entrada [3, 3, 3, 4, 4]

AMENAZA	DF	OP	CI	OP	FT	OP	ANA	OP	I	OP	NR
ENTRADA	3		3		3		4		4		
REGLA 71	b		b		b		b		b		b
GRADO DE PERTENENCIA	0,75	AND	0,75	AND	0,75	AND	1	AND	1	THEN (min)	0,75
AMENAZA	DF	OP	CI	OP	FT	OP	ANA	OP	I	OP	NR
ENTRADA	3		3		3		4		4		
REGLA 72	b		b		i		b		b		b
GRADO DE PERTENENCIA	0,75	AND	0,75	AND	0,25	AND	1	AND	1	THEN (min)	0,25
AMENAZA	DF	OP	CI	OP	FT	OP	ANA	OP	I	OP	NR
ENTRADA	3		3		3		4		4		
REGLA 73	b		b		b		i		b		b
GRADO DE PERTENENCIA	0,75	AND	0,75	AND	0,75	AND	0	AND	1	THEN (min)	0
AMENAZA	DF	OP	CI	OP	FT	OP	ANA	OP	I	OP	NR
ENTRADA	3		3		3		4		4		
REGLA 77	b		b		b		m		b		b
GRADO DE PERTENENCIA	0,75	AND	0,75	AND	0,75	AND	0	AND	1	THEN (min)	0
AMENAZA	DF	OP	CI	OP	FT	OP	ANA	OP	I	OP	NR
ENTRADA	3		3		3		4		4		
REGLA 88	b		b		i		i		b		b
GRADO DE PERTENENCIA	0,75	AND	0,75	AND	0,25	AND	0	AND	1	THEN (min)	0
AMENAZA	DF	OP	CI	OP	FT	OP	ANA	OP	I	OP	NR
ENTRADA	3		3		3		4		4		
REGLA 93	b		b		i		m		b		b
GRADO DE PERTENENCIA	0,75	AND	0,75	AND	0,25	AND	0	AND	1	THEN (min)	0

A continuación se procede a analizar la Regla 71, para el vector de entrada dado, obteniendo los siguientes valores:

ENTRADA: Daño Físico, VALOR: 3 (Bajo), se aplica la ecuación , lo cual arroja un resultado de 0,75.

El procedimiento anterior se aplica de igual forma a

las entradas Compromiso de la Información (CI) y Fallas Técnicas (FT).

ENTRADA: Acceso No Autorizado, VALOR: 4 (Bajo), se aplica la ecuación o , lo cual arroja un resultado de 1.

El procedimiento anterior se aplica de igual forma a la entrada Impacto (I).

Según las definiciones previas del modelo, para las cinco entradas se aplica el conector lógico AND y el valor mínimo de ellos se asigna a la SALIDA, la cual se definió previamente como el NIVEL DE RIESGO (NR); por lo tanto la salida obtenida es 0,75, tal y como aparece en la Fig.7.

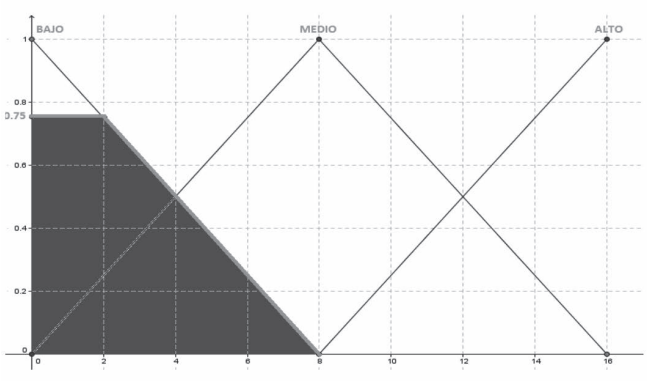

Fig.7. Salida entregada por la Regla 71

El análisis de la Regla 72 es casi igual a la regla anterior, por lo tanto sólo se analiza la entrada FT, que es la única que cambia.

ENTRADA: Fallas Técnicas, VALOR: 3 (Insignificante), se aplica la ecuación , lo cual arroja un resultado de 0,25.

Según las definiciones previas del modelo, para las cinco entradas se aplica el conector lógico AND y el valor mínimo de ellos se asigna a la SALIDA, la cual se definió previamente como el NIVEL DE RIESGO (NR); por lo tanto la salida obtenida es 0,25, tal y como aparece en la Fig.8.

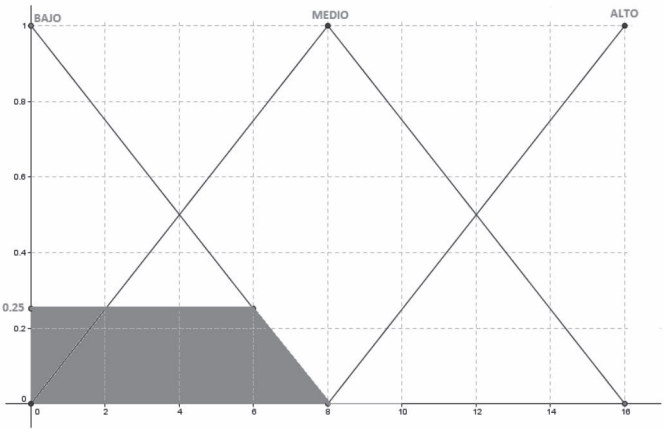

Fig.8. Salida entregada por la Regla 72

A continuación se procede a analizar la Regla 73, para el vector de entrada dado, obteniendo los siguientes valores: ENTRADA: Daño Físico, VALOR: 3 (Bajo), se aplica la ecuación , lo cual arroja un resultado de 0,75.

El procedimiento anterior se aplica de igual forma a la entrada Compromiso de la Información (CI) y Fallas Técnicas (FT).

ENTRADA: Acceso No Autorizado, VALOR: 4 (Insignificante), se aplica la ecuación , lo cual arroja un resultado de 0.

ENTRADA: Impacto (I), VALOR: 4 (Bajo), se aplica la ecuación ó , lo cual arroja un resultado de 1.

Según las definiciones previas del modelo, para las cinco entradas se aplica el conector lógico AND y el valor mínimo de ellos se asigna a la SALIDA, la cual se definió previamente como el NIVEL DE RIESGO (NR); por lo tanto la salida obtenida es 0.

Las salidas para las reglas 77, 88 y 93 tienen el mismo comportamiento de la regla 73.

Ahora se realiza la superposición de las ilustraciones resultantes de las reglas que fueron estimuladas (71, 72, 73, 77, 88 y 93), obteniéndose una ilustración final a la que se le aplica el método de desfuzzyficación seleccionado (centroide), obteniendo de esta manera el NIVEL DE RIESGO para el activo servidores. La Fig.9 muestra el resultado de la superposición de las Fig.7 y Fig.8

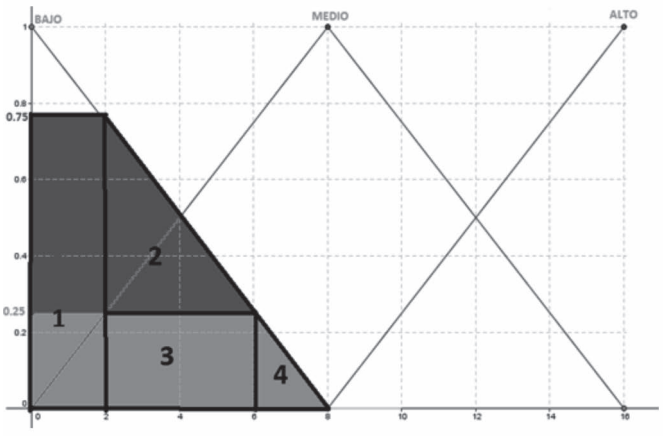

Fig 9. Superposición de las Fig.7 y Fig.8

Para este caso en particular, al aplicarle al prototipo el vector de entrada [3, 3, 3, 4, 4] se obtiene un NIVEL DE RIESGO de 2,75514, ver Fig.6.

Dado que el método de desfuzzyficación seleccionado fue el de CENTROIDE, a partir de la Figura 9 se procede a calcular de forma manual el centroide del polígono descrito en la citada ilustración, para lo cual se procede de la siguiente manera:

Cálculo de las áreas 1 a 4.
Área1= 3/2
Área2= 1
Área3= 1
Área4= 1/4
Cálculo del centroide de los polígonos 1 a 4.
Centroide1 (Rectángulo), $x1 = 1$

Centroide2 (Triángulo), $x2 = 10/3$
Centroide3 (Rectángulo), $x3 = 4$
Centroide4 (Triángulo), $x4 = 20/3$
Cálculo del centroide total

$$\dot{X} = (\sum A * X_n) / \sum A)$$

$$\text{Centroide} = [(3/2) * 1 + 1 * (10/3) + 1 * 4 + (1/4) * (20/3)] \; / \; [(3/2) + 1 + 1 + (1/4)]$$

Centroide = 2,8

D. Evaluación y Análisis de Resultados

Si se compara el resultado entregado por el prototipo con el resultado obtenido de forma manual, para la prueba UNO, se puede observar que son prácticamente iguales 2,75514 ≈ 2,8.

La Fig.10 muestra el grado de pertenencia de la variable de salida (Nivel de Riesgo) para el vector de entrada [3, 3, 3, 4, 4], aquí se observa que para un valor de 2,75514 el grado de pertenencia a la etiqueta lingüística BAJO es del 66%, mientras que para la etiqueta lingüística MEDIO es del 34%; por lo tanto, para este vector de entrada, el Nivel de Riesgo es BAJO, lo cual significa que se requiere de acciones a largo plazo, según lo establecido en la Tabla III.

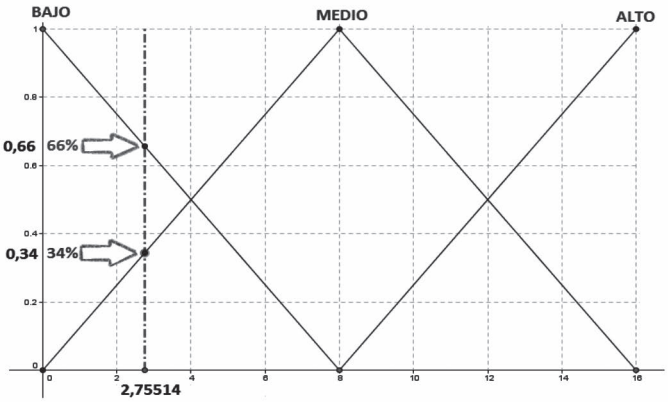

Fig.10. Grado de pertenencia de la variable Nivel de Riesgo para el vector [3, 3, 3, 4, 4]

VIII. Revisión de Trabajos Futuros

Se plantea como trabajo a futuro, evolucionar en el proceso de captura de los datos obtenidos de los expertos para cada activo de información, que permita mejorar la calidad y fiabilidad de los datos obtenidos en el modelo.

Integrar todas las reglas resultantes al modelo diseñado, de tal manera que permita la completa funcionalidad de éste.

Diseñar una plataforma computacional con una interfaz más amigable para el usuario, de tal manera que el modelo pueda ser utilizado e implementado como un mecanismo eficiente en el análisis de riesgos de la seguridad de la información y su utilización pueda ser promovida en diferentes ámbitos empresariales, facilitando la toma de decisiones.

Investigar sobre el uso de otras técnicas como algoritmos genéticos o redes neuronales aplicadas al análisis y evaluación de riesgos de seguridad de la información, que permitan la inclusión de nuevos parámetros o eliminar restricciones del sistema.

Con el fin de desarrollar las mejoras al modelo, aquí planteadas, es importante utilizar como referencia el material desarrollado en la Universidad Politécnica de Madrid, particularmente por los Ingenieros Alfonso Mateos Caballero, Antonio Jiménez Martín y Eloy Vicente Cestero; uno de cuyos trabajos hace parte de la referencias de este documento, [7].

IX. Conclusiones

La teoría de la lógica difusa aplicada para realizar el análisis y evaluación de riesgos de seguridad en los activos de información, genera y entrega datos más exactos de nivel de riesgo, que utilizando la metodología convencional cualitativa. El modelo brinda al usuario la posibilidad de una mayor interpretación a la subjetividad envuelta en el análisis.

Una vez analizados los resultados entregados por el prototipo se puede aceptar la hipótesis inicial de este proyecto, donde se afirma que aplicando lógica difusa se puede generar un modelo para la evaluación del análisis de riesgos de seguridad de la información.

Matlab como herramienta de desarrollo y sobre todo su toolbox de lógica difusa, permite la creación de un controlador difuso de una manera rápida, intuitiva, estructurada, sencilla y de fácil entendimiento.

La lógica difusa se perfila como una alternativa importante para el desarrollo de modelos y sistemas basados en el conocimiento, que constituyen una verdadera herramienta de apoyo a la toma de decisiones empresariales en lo referente a análisis y gestión de riesgos.

Referencias

[1] R. Puello. "Procesado y optimización de espectros RAMAN mediante técnicas de lógica difusa: Aplicación a la identificación de materiales pictóricos". Universidad Politécnica de Cataluña. 2005.

[2] Proyecto de Norma Técnica Colombiana NTC-ISO 27005. Bogotá D.C.: ICONTEC, 2008. 96p.

[3] A. Angarita y C. Tabares. Análisis de riesgos para el proceso administrativo: Departamento de Informática en la Empresa de Acueducto y Alcantarillado de Pereira S.A. E.S.P., basados en la Norma ISO 27005. Trabajo de grado Especialista en Redes de Datos. Pereira: Universidad Tecnológica de Pereira. Facultad de Ingenierías Eléctrica, Electrónica, Física y Ciencias de la Computación, 2012. 104 p.

[4] J. Burgos y P. Campos. Modelo Para Seguridad de la Información en TIC. 2008, versión 2. Available from Internet: http://ceur-ws.org/Vol-488/paper13.pdf
[5] J. Ramió. Libro Electrónico de Seguridad Informática y Criptografía. Madrid, España. 2006. 1106 diapositivas.
[6] P.Ponce. Inteligencia Artificial con aplicaciones a la ingeniería. 1 ed. México D.F.: Editorial Alfaomega, 2011. 348 p. ISBN 978-84-267-1706-1.
[7] E. Cestero. Un enfoque borroso para el análisis y la gestión de riesgos en sistemas de información. Trabajo de Grado Máster Universitario en Inteligencia Artificial. Madrid: Universidad Politécnica de Madrid. Facultad de Informática, 2013.106 p.
[8] T. Feagans and W. Biller. Fuzzy Concepts in the Analysis of Public Health Risks. Springer US. 1980. 391-404 p.
[9] A. Markowski and S. Mannan. Fuzzy Risk Matrix. Journal of Hazardous Materials, Volume 159, Issue 1, 15 November 2008. 152-157 p.
[10] B. Martín del Brío y A. Sanz. Redes neuronales y sistemas borrosos. 3 ed. España: Ra-Ma Editorial, 2006. 442 p. ISBN 978-84-7897-743-7.
[11] A. Benavides. Diseño y aplicaciones de un sistema de priorización de riesgos basado en lógica difusa y proceso analítico jerárquico. Trabajo de grado Ingeniero Químico. Bucaramanga: Universidad Industrial de Santander. Facultad de Ingenierías Físico-Químicas, 2010.79 p.
[12] S.N. Sivanandam, S. Sumathi and S.N. Deepa. Introduction to Fuzzy Logic using Matlab. Editorial Springer, 2007. 441 p.
[13] C. Alberts and A. Dorofee. Managing Information Security Risks: The OCTAVE Approach. Addison-Wesley, 2002. 512 p. ISBN 0321118863.
[14] S-M. Chen. New Methods for subjective mental workload assessment and fuzzy risk analysis. En: Cybernetics and Systems: An International Journal, Volume 27, Issue 5, 1996.

Jorge Iván Ríos Patiño. Profesor Titular del Programa Ing. Sistemas y Computación – Universidad Tecnológica de Pereira. Es Ing. Industrial – Universidad Tecnológica de Pereira, MSc Informática e Ing. Del Conocimiento – Universidad Politécnica de Madrid y PhD (c) Informática – Universidad Politécnica de Madrid.

Es director de la Maestría en Ingeniería de Sistemas y Computación de la Universidad Tecnológica de Pereira desde junio de 2009. Sus áreas de actuación son la Inteligencia Artificial, Ciencias de la Computación y de la Información.

Alexis Armando Angarita Vivas. Ingeniero de Sistemas y Computación – Universidad Tecnológica de Pereira en el 2003, Especialista en Redes de Datos – Universidad Tecnológica de Pereira en el 2013 y MSc (c) Ingeniería de Sistemas y Computación – Universidad Tecnológica de Pereira.

Actualmente trabaja en la división de sistemas de la empresa Aguas y Aguas de Pereira.

Sus campos de interés son el diseño de redes de comunicaciones, seguridad de la información y la gestión de tecnología.

César Augusto Tabares Isaza. Ingeniero Electricista – Universidad Tecnológica de Pereira en el 1988, Especialista en Redes de Datos – Universidad Tecnológica de Pereira en el 2013 y MSc (c) Ingeniería de Sistemas y Computación – Universidad Tecnológica de Pereira.

Actualmente trabaja en la Secretaría de Educación Municipal de Pereira como docente en el área de matemáticas, y adicionalmente se desempeña como catedrático auxiliar en la Universidad Tecnológica de Pereira.

Sus campos de interés son el diseño de redes de comunicaciones, seguridad de la información y la gestión de tecnología.