UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS

FACULTAD 2

MODELO PARA LA EVALUACIÓN DEL RIESGO DE SEGURIDAD DE LA INFORMACIÓN EN LOS SISTEMAS GESTORES DE BASES DE DATOS

Tesis presentada en opción al grado científico de Doctor en Ciencias Técnicas

YASSER AZÁN BASALLO

La Habana, 2017

UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS
FACULTAD 2

MODELO PARA LA EVALUACIÓN DEL RIESGO DE SEGURIDAD DE LA INFORMACIÓN EN LOS SISTEMAS GESTORES DE BASES DE DATOS

Tesis presentada en opción al grado científico deDoctor en Ciencias Técnicas

Autor:MSc. Yasser Azán Basallo

Tutores: Dra.C. Vivian Estrada Sentí,Prof. Titular

Dra.C.Natalia Martínez Sánchez,Prof. Titular

La Habana, 2017

Agradecimientos

A la Revolución, por darme esta oportunidad.

A mis tutoraspor su apoyo.

A mis compañeros del equipo AuditDB, por el trabajo compartido.

A Salvador yAlberto, por sus consejos y aliento.

A mi novia Ivonne Guerra Anaya quien sufrió a mi lado todas las adversidades.

A todos los que de una forma u otra contribuyeron con mi formación científica.

Dedicatoria

A mi familia:

En especial a mi mamá y a mi papá que mucho han sacrificado por mí para que alcanzara esta meta.

SÍNTESIS

- En las auditorías de seguridad informática se realizan las evaluaciones de riesgo de seguridad de la información a sistemas de cómputo. En un estudio realizado se evidenció que existen diferentes niveles de experticia entre los auditores, lo que provoca posibles diferencias entre la evaluación real del riesgo y la calculada por el experto. La evaluación del riesgo se clasifica según su impacto en: Alto, Medio o Bajo por lo que se pueden generar ambigüedades en el resultado de la evaluación. Las listas de chequeo de seguridad, tienen una fuerte dependencia de la presencia del auditor en base de datos para el análisis del riesgo. En aras de facilitar el trabajo de los auditores, se propone un modelo basado en el conocimiento y la lógica difusa para la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos. El modelo se aplica en el Departamento de Seguridad Informática de la entidad ETECSA, la cual tiene entre sus principales responsabilidades, garantizar y mantener la integridad de los datos en los sistemas gestores de bases de datos y aplicaciones web que soportan el trabajo de las telecomunicaciones en Cuba. Los resultadosobtenidosmuestran que se aprovecha la experiencia acumulada en las auditorías anteriores de este tipo y se mejora la exactitud de los resultados en la evaluación del riesgo de seguridad de la información.

ÍNDICE

INTRODUCCIÓN

Los avances en los sistemas de información (SI) y las tecnologías originan grandes resultados para organizaciones, negocios y otras agencias en términos de productividad del trabajo, almacenamiento de la información, administración y oportunidad de ventajas competitivas. Mientras los SI ofrecen extraordinarios beneficios, también representan mayores niveles de riesgo de modo significativo y sin precedentes, para las operaciones organizacionales.Los negocios, hospitales, escuelas, universidades, agencias gubernamentales y bancos dependen fuertemente de los SI; esto incrementa la necesidad de la seguridad de la información(Quigley, M., 2008).

Los gestores de bases de datos son uno de los SI con frecuentes ataques a las vulnerabilidades existentes en las mismas(Ramakanth, D. y Vinod, K., 2011): El Departamento de Justicia de los Estados Unidos acusó a un ciudadano por el robo de 130 millones en tarjetas de crédito usando ataques de inyección de SQL. Aproximadamente 500.000 páginas web que usaban como servidor el Microsoft IIS y el servidor de SQL, fueron atacadas entre abril y agosto del 2008 usando la inyección de SQL.

Todos los años se han reportados un número conocido de vulnerabilidades reportadas al aplicar la inyección de SQLpor el El Instituto Nacional de Normas y Tecnologíade los Estados Unidos de América (por las siglas en inglés: NIST)("N.I.S.T.", 2017a): en el 2010, 2011se registraron 4,639 y4,150de ataques por vulnerabilidad respectivamente, mientras que en el 2015 y 2016 fueron 6,488 y 6,449respectivamente.

Este mismo instituto señala que muchos otros tipos de vulnerabilidades de bases de datos se han acrecentado en años recientes. Los siguientes datos de la vulnerabilidaddenominada como:Permisos, Privilegios, y Control de Accesoes muestra de la afirmación("N.I.S.T.", 2017b): En el 2012 y 2013 se registraron 5,288 y 5,186 de ataques por vulnerabilidad

respectivamente, mientras que en el 2014, 2015 y 2016 fueron 7,937, 6,488 y 6,449respectivamente.

A los anteriores datos de ataques a las vulnerabilidades, se conoce que el 96% de los datos sustraídos durante 2012, provenían de bases de datos (Quisbert, A. E. V., 2014).

Lo expresado anteriormente, refleja la importancia de proteger los datos ante las vulnerabilidades detectadas en los sistemas gestores de bases de datos, donde la seguridad de la información tiene como fin la protección de los datos y el control del acceso, uso, divulgación, interrupción o destrucción no autorizada de los SI("A.E.C.", 2013).

Para la seguridad de la información, son fundamentaleslos controles de seguridad a los sistemas informáticos. Entre las razones se encuentran: el impacto de los controles de seguridad de los sistemas de información a otros controles generales, la vulnerabilidad de los sistemas de computadoras hacia la pérdida de recursos, el impacto del fracaso de la seguridad en datos confiables, la posibilidad de faltar a los cumplimientos con los requisitos legales, la posibilidad de pérdida de la contingencia si el riesgo del proceso de los datos son graves y no están asegurados, la vulnerabilidad de los sistemas de computadora a ser usados sin autorización, entre otras.

Uno de los pasos para garantizar la seguridad de los datos es realizarauditorías de seguridad a los sistemas computacionales. El mismo es un proceso relevante para la seguridad de la información. El autor de esta investigación comparte el concepto de auditorías de seguridad a los sistemas computacionalesformulado por Muñoz como: la revisión exhaustiva, técnica y especializada que se realiza a todo lo relacionado con la seguridad de un sistema de cómputo, sus áreas y personal, así como a las actividades, funciones y acciones preventivas y correctivas que contribuyan a salvaguardar la seguridad de los equipos computacionales, las bases de datos, redes, instalaciones y usuarios del sistema. Es también la revisión de los planes de contingencia y medidas de protección para

la información, los usuarios y los propios sistemas computacionales y en sí para todos aquellos aspectos que contribuyan a la protección y salvaguarda en el buen funcionamiento del área de sistematización, sistemas de redes o computadoras personales. Se incluye la prevención y erradicación de los virus informáticos(Muñoz, C. R., 2002).

En las auditorías de seguridad a los sistemas computacionalesse realiza la evaluación del riesgo de seguridad de la información. La evaluación del riesgo es el proceso de identificación de las amenazas a los sistemas de información, la determinación de la probabilidad de ocurrencia de la amenaza y la identificación de las vulnerabilidades del sistema que podrían ser explotadas por la amenaza (Ms-Isac, M.-S. I. S. a. a. C., 2010).

En el mundo existe un conjunto de metodologías como son: MAGERIT, COBRA, BPIRM, NIST SP800-30, IT-Grundschutz, MEHARI, EBIOS, ISRAM, CRAMM, OCTAVE,CORAS, entre otras, con el objetivo de indicar como evaluar el riesgo de este tipo.

Los métodos están clasificados principalmente en cuantitativos y cualitativos.Los métodos cuantitativos de análisis de riesgos utilizan herramientas matemáticas y estadísticas para representar el riesgo(Tóth, G. N. y Berek, L., 2010). Un análisis cuantitativo proporciona una medida de la magnitud de los impactos, que se puede utilizar en el análisis costo-beneficio de los controles recomendados. La desventaja es que al depender de los rangos numéricos utilizados para expresar la medición, el significado del análisis cuantitativo de impacto puede ser poco claro y requerirde una interpretación del resultado de una manera cualitativa (según criterios de los especialistas). Por ejemplo en: Alto, Medio o Bajo.

Los métodos cualitativos de análisis de riesgo son realizados con la ayuda de los adjetivos en lugar de usar las matemáticas(Karabacak, B. y Sogukpinar, I., 2005). La principal ventaja del análisis del impacto cualitativo es que se le da prioridad a los riesgos e identifica las áreas de mejora. La desventaja del análisis cualitativo es que no proporcionan

mediciones cuantificables específicas de la magnitud de los impactos, por lo tanto, hacer un análisis de costo-beneficio de los controles recomendados es difícil. El resultado es muy dependiente de las ideas de personas que llevan a cabo el análisis del riesgo(Karabacak, B. y Sogukpinar, I., 2005).

En Cuba se realizan auditorías de forma periódica a los sistemas computacionales y se estima el riesgo de seguridad de la información. Entre las entidades que realizan auditorías de forma periódica se encuentra la empresa ETECSA, a la que pertenece el Departamento de Seguridad Informática (DSI) que contempla entre sus principales responsabilidades la de garantizar y mantener la integridad de los sistemas gestores de bases de datos (SGBD), sistemas operativos y aplicaciones web que soportan el trabajo de las telecomunicaciones en Cuba. Esta empresa constituye el caso de estudio de la presente investigación.

ETECSA toma como base para para la auditoría de seguridad informática la experiencia de la capacitación dada por la empresa Ernest& Young (EY), la cual brinda consultoría y asesoría en tema entre otros de seguridad. En el 2016 tenía registrado 212.000 empleados, en ventas $28.7 000 millones("Forbes", 2017) y con presencia en más de 150 países ("E.Y.", 2017).En ETECSA, para evaluar el riesgo de seguridad de la información (RSI)en los SGBD, por lo general se emplean:

- Aplicaciones para monitorear la configuración de seguridad de la información (Ejemplos: Secure Oracle Auditor, Secure SQL Auditor, Microsoft Baseline Security Analyzer, DB Audit y otros).
- Las listas de chequeode seguridad o matrices de diagnóstico del Centro para la Seguridad de Internet.

Para realizar las auditorías de seguridad informáticas, los especialistas se basan en la técnica: listas de chequeo, para comprobar la correcta configuración de cada uno de los parámetros. A partir de las listas de chequeo de seguridad, se detectan las configuraciones

de seguridad puestas en el servidor y el experto valora el nivel de RSI, centrado en laguíaproporcionada por las mismas.

Las listas de chequeo son usadas para facilitar la recolección de información pertinente. Estas pueden tener muchas formas. Pueden ser simples listas de preguntas de si o no, o de preguntas de final abierto yprecisarrespuestas en formas de redacción. Pueden ser cortas y con enfoques limitados en la específica operación o actividad en cuestión; o pueden ser extensas en alcance y cubrir lo común en lo concerniente a la seguridad de todas las operaciones de la compañía(Broder, J. F. y Tucker, G., 2011).

Una vez recogidos los datos en el monitoreo de los SGBD a través de la lista de chequeo, se realiza la evaluación de los resultados. Este proceso se efectúa para determinar la evaluación del riesgo, la cual se expresa en los términos: Alto, Medio y Bajo.

A partir de los estudios realizados, el autor de esta investigación ha identificado la existencia de dificultades o limitaciones relacionadas con la evaluación del riesgo de seguridad de la información tales como:

- El criterio de losespecialistas en las auditorías,tiene alta influencia en el resultado final en la evaluación del RSI.
- Existen diferentes niveles de experticia entre los auditores para evaluar los SGBD con las listas de chequeo, lo que provoca diferencias entre la evaluación real del riesgo y la estimada por el auditor (la subjetividad incide negativamente en los niveles de exactitud).
- El 100% de los auditores entrevistados manifestaron que sería muy provechoso contar con información sobre situaciones pasadas y las decisiones tomadas para así mejorar sus propias decisiones.

- Las listas de chequeo tienen una fuerte dependencia de la opinión del auditor en el resultado del análisis del RSI en los SGBD, lo que provoca diversidad de opiniones antes situaciones iguales o semejantes(Piattini, M. G. V. y De Peso, E. N., 2001).
- La evaluación del RSI en los SGBD se expresa en los términos: Alto, Medio o Bajo, por lo que para cada auditor, constituye una medida ambigua, sin límites precisos.
- Las metodologías de evaluación del riesgo de seguridad para los sistemas de información son generales, por lo que es complejo evaluar el RSI a sistemas de cómputo en entidades con diferentes características de seguridad.
- Para proporcionar el resultado de la evaluación del RSI, los auditores se pueden tardar horas o días, por lo que existe demora en conocer los resultados en la auditoría de seguridad informática y por tanto también en la toma de decisiones.

Por otra parte, existen técnicas de la Inteligencia Artificial (IA) (Bello, R., 2002), (Guida, G. y Tasso, C., 1994) como los sistemas basados en el conocimiento (Hand, D. J., 1997), (Guida, G. y Tasso, C., 1994) válidas para automatizar el proceso de evaluación del riesgo de seguridad de la información en los SGBD dado sus aspectos afines.

Los sistemas basados en el conocimiento utilizan conocimiento sobre un dominio específico y la solución que se obtiene es similar a la dada por una persona experimentada en el dominio del problema. En el caso objeto de estudio la experiencia de los auditores es fundamental para la toma de decisiones en los RSI. La incertidumbre, imprecisión y/o vaguedad están presentes en casi todos los datos a utilizar en el proceso de evaluación del riesgo de seguridad de la información en los SGBD.

Diferentes tipos de conocimiento dan lugar a diferentes tipos de sistemas basados en el conocimiento, entre ellos los sistemas basados en reglas (Bello, R., 2002), los sistemas

basados en probabilidades (Minka, T. P., 2001), (Lerner, U., 2002), sistemas expertos conexionistas o redes expertas (Ohno-Machado, L., 1996), (Dutta, S. y Bonissone, P., 1991) y los sistemas basados en casos (Kolodner, J. L., 1992),(Gutiérrez, I. M. y Bello, R. E. P., 2003).

No todos los paradigmas para crear sistemas basados en el conocimiento facilitan la concepción de un sistema inteligente para evaluar el riesgo de seguridad de la información en un SGBD, donde lo fundamental para su desarrollo es contar con la información acumulada de auditorías realizadas y el manejo de la incertidumbre en los datos con los cuales se trabaja. Por lo que las características distintivas de los sistemas basados en casos, son factores a estudiar para concebir este tipo de sistema inteligente de apoyo a dicho proceso.

Por lo antes planteado se identificóel siguiente**problema científico**: ¿cómo aumentar la exactitud y disminuir el tiempo de respuesta en la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos?

El **objeto de estudio**lo constituye la evaluación del riesgo de seguridad para los sistemas computacionales.

El **campo de acción:** los procesos de evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos.

Para darle una solución efectiva al problema, se plantea como **objetivo general**: desarrollar un modelo basado en el conocimiento y la lógica difusa para aumentar la exactitud y disminuir el tiempo de respuesta en la evaluación del RSI en los SGBD.

Para alcanzar el objetivo general, se plantean los siguientes**objetivos específicos:**

1. Construir el marco teórico conceptual de la investigación relacionada con la evaluación del RSI para los SGBD.

2. Elaborar un modelo basado en el conocimiento y la lógica difusa que contribuya a aumentar la exactitud y la disminucióndel tiempo de respuesta de la evaluación del RSI en los SGBD.
3. Realizar la implementación computacional del modelo propuesto.
4. Validar el modelo propuesto.

Después de haber analizado la literatura para conformar el marco teórico, se formula la siguiente**hipótesis de investigación:** un modelo para la evaluación del RSI en los SGBD, basado en el conocimiento y la lógica difusa logra aumentar la exactitud y disminuir el tiempo de respuesta de este proceso.

Para dar cumplimiento al objetivo propuesto, se han utilizado diversos métodos científicosy procedimientos teóricos y empíricos, en la búsqueda y procesamiento de la información. Los fundamentales son:

Métodos teóricos

1. Análisis-síntesis: para el estudio de las fuentes bibliográficas existentes referente al tema, con el cual se identificaron los elementos más importantes y necesarios para dar solución al problema planteado.
2. Hipotético-deductivo: permitió definir la hipótesis de la investigación y proponer nuevas líneas de trabajo relacionadas con la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos.
3. Histórico-lógico: permitió el estudio del comportamiento y evolución de las diferentes posiciones respecto a las diferentes herramientas de evaluación del riesgo de seguridad de la información para los sistemas de cómputo.
4. El método sistémico es empleado en la concepción del modelo propuesto y para el desarrollo de las herramientas de software y lograr que los elementos que conforman las mismas constituyan un todo que funcione de manera armónica.

5. La modelación: Para la concepción del modelo propuesto y la descripción del funcionamiento de los componentes por los cuales está constituido el mismo.

Métodos empíricos

1. Análisis documental: en la consulta de la literatura especializada en las temáticas.
2. Entrevista: se aplicó para obtener toda la información necesaria respecto a cómo es realizadoel proceso deevaluación del riesgo de seguridad de la información en las auditorías de seguridad informática a las Tecnologías de la Información en cuanto a los problemas comunes a los que se enfrentan los auditores de seguridad informática. Se elaboró una guía de preguntas, garantizando que fueran descritos los aspectos fundamentales.
3. Encuesta: para obtener un diagnósticodel proceso de evaluación del riesgo de seguridad de la información en las auditorías de seguridad informática a las Tecnologías de la Información en cuanto a las metodologías, estándares y técnicas empleadas, las aplicaciones informáticas utilizadas y los problemas comunes a los que se enfrentan los auditores de seguridad informática. Además como instrumento en el proceso de validación en las distintas etapas concebidas.
4. Método de consulta a expertos: para la validación teórica de los componentes y la integralidad del modelo propuesto.
5. El estudio de casos y el experimento: para evaluar en la práctica la contribución del modelo en la mejora de la exactitud y en la disminución del tiempo de respuestade la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos.
6. Se aplica la Técnica de Iadov, para el estudio de la satisfacción de los usuarios con las herramientas (SASGBD y HCRA) desarrolladas.

La novedad científica del trabajo se expresa en los siguientes aportes teóricos y prácticos:

- **Aportes teóricos:**
  - ✓ La fundamentación y sistematización de un modelo basado en el conocimiento y la lógica difusapara la evaluación del riesgo de seguridad de la informaciónen los SGBD.
  - ✓ Cinco algoritmos que permiten la evaluación del riesgo y agilizar el proceso de auditoría de los SGBD respecto a la seguridad de la información.
- **Aportes prácticos:**
  - ✓ Una solución informática compuesta por la aplicación HCRA y el SASGBD como instancias de la extensión del modelo propuesto, para la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos.
  - ✓ Un conjunto de indicaciones metodológicas para el empleo del modelo propuesto.

Estructura del documento

El presente documento se encuentra estructurado en tres capítulos. En el primero de ellos setiene como objetivo analizar los diferentes aspectos teóricos relacionados con la evaluación del riesgo de seguridad de la información. Se analizan y adoptan los principales conceptos que facilitan el estudio y compresión de la temática.

En el segundo capítulo se describen los componentes que forman parte del modelo para la evaluación del riesgo de seguridad de la información. Se describen las partes fundamentales de cada componente, así como el funcionamiento del modelo y las indicaciones metodológicas del mismo.

En el tercer capítulo se describe el proceso para la validación del modeloa través de los análisis pertinentes en cada una de las actividades que lo integran. El capítulo concluye con una triangulación metodológica de los resultados.

Finalmente se presentan las conclusiones y recomendaciones derivadas de la investigación, las fuentes bibliográficas consultadas y los anexos que apoyan la comprensión y dan información adicional sobre el trabajo realizado.

CAPÍTULO 1.

MARCO TEÓRICO REFERENCIAL DE EVALUACIÓN DEL RIESGO DE SEGURIDAD DE LA INFORMACIÓN

CAPÍTULO 1.MARCO TEÓRICO REFERENCIAL DEEVALUACIÓN DEL RIESGO DE SEGURIDAD DE LA INFORMACIÓN

En el presente capítulo se abordan los principales referentes teóricos y estado del arte que fundamentan la investigación. Se presenta una valoración de los principales procedimientos conocidos a nivel internacionalrelacionados con la evaluación del riesgo de la seguridad de la información.

1.1 Mecanismos para la estimación del riesgo de seguridad de la información

En la actualidad existen varios mecanismos para realizar la evaluación del riesgo de seguridad de la información de las Tecnologías de la Información (TI). En los siguientes subepígrafes se muestran una representación de los mismos.

1.1.1 Metodologías de análisis del riesgo de seguridad de la información

Para la evaluación del riesgo de seguridad de la información en los sistemas de cómputo existen en la actualidad un gran número de metodologías enfocadas a este tema. Básicamente hay dos tipos de métodos de análisis del riesgo (Tóth, G. N. y Berek, L., 2010):los métodos cualitativos, los cualesson realizados con la ayuda de los adjetivos en lugar de usarlas matemáticas.Los métodos cuantitativos de análisis de riesgos, los cuales utilizan herramientas matemáticas y estadísticas para representar el riesgo. Por último están las metodologías híbridas o mixtas, las cuales incluyen las anteriores.

Cualitativas

La principal ventajadel análisis del impactocualitativoes que se le da prioridad alos riesgose identificaáreas de mejorainmediatamentepara hacer frente alasvulnerabilidades, según afirma la Guía de Administración de Riesgo (GAR) del Instituto Nacional de Estándares y Tecnología (conocida por las siglas en inglés NIST) (Stoneburner, Goguen, G., y Feringa, A., 2002).

Varias de estas metodologías se basan en las diversas opiniones de los cuestionarios o de las encuestas donde los participantes pueden estar o no muy consciente de los recientes desarrollos en el que se trate el área o de reuniones o talleres (Pandey, S. K., 2012). Por lo general cumplen con los estándares de administración del riesgo: ISO/IEC 27001, ISO/IEC 15408, ISO/IEC 17799, ISO/IEC 13335 y la ISO/IEC 21827. El objetivo de estas metodologías es priorizar los riesgos, identificar la lista de amenazas y las contramedidas pertinentes requeridas en un nivel relativamente técnico. Ejemplo de estas metodologías son: CORAS, COBRA, EBIOS, IT-Grundschutz (IT BaselineProtection Manual), BPIRM("E.N.I.S.A.", 2013; Coles, R. S. y Moulton, R., 2003; Dimitrakos, T., Raptis, D., Ritchie, B., y Stølen, K., 2002; Hisham, M. H. y Brunil, D. R., 2009; Karabacak, B. y Sogukpinar, I., 2005; Kouns, J. y Minoli, D., 2011; Nikolić, B. y Ružić-Dimitrijević, L., 2009; Pandey, S. K., 2012; Peltier, T. R., 2001; Sikianakis, E. C., Antonakis, N., y Stolen, K., 2003; Tóth, G. N. y Berek, L., 2010), entre otras.

El análisis cualitativono proporcionamedicionescuantificablesespecíficas dela magnitudde los impactos, por lo tanto,hacerun análisis de costo-beneficio de los controlesrecomendados son difíciles(Stoneburner y otros, 2002). Los métodos cualitativos son inconvenientes debido a su naturaleza, las cuales rindenresultados inconsistentes(Karabacak, B. y Sogukpinar, I., 2005).Debido a quelos métodos cualitativosno utilizanherramientas como las matemáticasy las estadísticas para modelarel riesgo, el resultado delmétodo esvastamentedependiente delas ideas de personasque llevan a caboel análisis de riesgo.Esto provoca que sea muy dependiente de la participación del experto (Carvalho, F. D. y Silva, E. M. D., 2006).

Cuantitativas

Las metodologías cuantitativas asignan valores a los riesgos individuales y los combinan matemáticamente para quedemuestren sus efectos relativos o absolutos.La necesidad de realizar la medición del riesgo es la posibilidad de cuantificar el costo potencial y la expresión lógica de la frecuencia de ocurrencia(Broder, J. F. y Tucker, G., 2011). Estos métodos usan métodos estadísticos y herramientas matemáticas para estimar el riesgo (Tóth, G. N. y Berek, L., 2010).En este grupo de metodologías están: ISRAM e ISAMM ("E.N.I.S.A.", 2013; "N.I.S.T.", 2011; Karabacak, B. y Sogukpinar, I., 2005), entre otras.

Las metodologías cuantitativas tienen las ventajas de ser objetivas e independientes. Brinda una base sólida para el análisis costo-beneficio de las salvaguardas y credibilidad para auditar y gestionar, especialmente en gestión corporativa(Areitio, J. B., 2008). La principal ventaja deun análisiscuantitativo de impactoes que proporcionauna medida de lamagnitudde los impactos, empleado en el análisis coste-beneficio de los controles recomendados(Tóth, G. N. y Berek, L., 2010).

Estas metodologías pueden tener identificación y evaluación del riesgo(Kouns, J. y Minoli, D., 2011) y la elaboración de medidas defensivas (Tóth, G. N. y Berek, L., 2010). Ademástener compatibilidad con la norma ISO 27001, lo que aumenta su aplicabilidad (Pandey, S. K., 2012).

Algunas de estas metodologías tienen capacidad para simular el efecto de la reducción en el riesgo de expectativa de pérdidaanual(ALE), en cada control y comparar esto con su costo de implementación ("E.N.I.S.A.", 2013).El problema con este tipo de metodologías es el esfuerzo para identificar los valores de los recursos (especialmente los valores de los recursos intangibles) y desarrollar válidos datos de frecuencias de amenazas (Peltier, T. R., 2001).

Otra dificultad de adoptar una metodología de análisis de riesgo cuantitativa es la complejidad de determinar el impacto de un evento no deseado y principalmente, la falta de datos suficiente para poder determinar de manera exacta las funciones de distribución de probabilidad para las amenazas más comunes(García, J. M., 2008). Añade que tiene los inconvenientes de no ser fiables para eventos desconocidos o impactos impensables.Consumen mucho tiempo y es costoso a la hora de hacerlo bien. Los números implicados son subjetivos. Es incapaz de discriminar entre eventos de amenaza baja e impacto alto y eventos de amenaza de bajo impacto y elevada frecuencia.

Una de las desventajas de las metodologías cuantitativas es que en dependencia de los rangos numéricos utilizados para expresar la medición, el significado del análisis cuantitativo de impacto puede ser poco claro y requerir una interpretación del resultado de una manera cualitativa(Tóth, G. N. y Berek, L., 2010). Otro punto en desventaja considerada es que las medidas de matemáticos intensivos utilizados para el riesgo del modelo de entornos complejos hacen que el proceso sea más difícil(Karabacak, B. y Sogukpinar, I., 2005).

Híbridas o mixtas

Es una selección combinada de las anteriores metodologías, que son generalmente aplicadas para implementar los componentes por las cuales están conformadas. Para ello se utilizan los fragmentos de información y las métricas a ser calculadas y recolectadas(Tóth, G. N. y Berek, L., 2010). Dentro de las metodologías híbridas conocidas están: OCTAVE,Magerit ySOMAP ("S.O.M.A.P.", 2007; Cole, E., 2011; Eloff, J. H., Labuschagne, L., y Badenhorst, K. P., 1993; Hisham, M. H. y Brunil, D. R., 2009; Pandey, S. K., 2012; Públicas, M. D. H. Y. A., 2012; Shoniregun, C. A., 2006; Smith, S. T., 1989; Whitman, M. y Mattord, H., 2011); (Jiankang, L. y Bing, J., 2012)y otras.

Como característica similar este tipo de metodologías pueden orientar una estrategia de protección de la organización y los planes de mitigación de riesgos para los más altos priorizados (Whitman, M. y Mattord, H., 2011).

En las revisiones hechas por el autor de la presente investigación a las metodologías cuantitativas, cualitativas o mixtas, no se encontró una propuesta de medición de la influencia en su conjunto, de todas las amenazas detectadas. Sino que solo se determina el riesgo de cada amenaza o del activo en forma cualitativa o cuantitativa, por lo que es otro elemento que influye para que las mismas necesiten la acción del experto para llegar a la evaluación del riesgo de seguridad de la información.

En el proceso de investigación se identificó que los especialistas de seguridad informática y auditores de ETECSA utilizan un procedimiento muy cercano a la técnica análisis mediante tablas, propuesta por la metodología Magerit en la versión 3.0 (M.H.A.P., 2012a).

1.1.2 Estándar y protocolo

ISO/IEC 27005

Es un estándar para proporcionar una guía sobre el proceso de administración del riesgo de seguridad de la información, necesaria para un efectivo sistema de administración de la seguridad de la información. Aunque este estándar es considerado para la administración del riesgo, una considerable parte trata sobre la evaluación del riesgo, la cual es una parte principal del programa de administración del riesgo. El estándar está alineado con la NIST 800-30 y escrita con un alto nivel de perspectiva (Talabis, M. y Martin, J., 2012).

El estándar tiene tres pasos en la sección que trata con la evaluación del riesgo: identificación de riesgo, estimación del riesgo y evaluación de riesgo. En esta última se proporciona los valores de la probabilidad de un incidente y la estimación del riesgo y se realiza una priorización de las mismas(Talabis, M. y Martin, J., 2012).

ISO 27005 derogólas normasISO/IEC TR 13335-3:1998 e ISO/IEC TR 13335-4:2000, y proporciona desde su publicación en Junio del año 2008, un conjunto de directrices para la correcta realización de un análisis delriesgo. La ISO 27005 no proporciona una metodología concreta de análisis delriesgo, sino que describe a través de su conjunto de cláusulas el proceso recomendado de análisis y las fases que lo conforman: establecimiento del contexto (Cláusula 7), evaluación del riesgo (Cláusula 8), tratamiento del riesgo (Cláusula 9), aceptación del riesgo (Cláusula 10), comunicación del riesgo (Cláusula 11), monitorización y revisión del riesgo (Cláusula 12)(Mendenhall, W., Beaver, R. J., y Beaver, B. M., 2010).

Security Context Automation Protocol (SCAP)

Este protocolo fue especificado por TheNationalInstituteofStandardsandTechnology, la misma consiste enun conjunto deespecificaciones paraestandarizarel formato ynomenclaturaen la queel software de seguridadcomunicainformación sobrelos defectos del softwarey configuracionesde seguridad.La estandarización dela información de seguridadfacilita laherramienta deinteroperabilidad ypermite obtener resultadospredeciblespor medio delSCAP entre software de seguridaddispares. ElPrograma de Validación deSCAPofrece a los proveedoresla oportunidad de teneruna verificación independientede queel software de seguridadprocesacorrectamente la información de seguridad del SCAPproporciona una salidanormalizada ("N.I.S.T.", 2013).

A pesar que a través de los componentes del SCAP se pueden chequear las configuraciones de seguridad y se realizan reporte de las mismas, este no es capaz de evaluar el riesgo de seguridad de la información.

1.1.3 Las listas de chequeo de seguridad para los SGBD

Las listas de chequeo son aplicadas para facilitar la recolección de información pertinente(Broder, J. F. y Tucker, G., 2011). Estas pueden tener muchas formas. Pueden

ser simples listas de preguntas de si o no, o de preguntas de final abierto con respuestas en formas de redacción. Ellas pueden ser breves y de enfoques limitados en la específica operación o actividad en cuestión, o pueden ser extensas en alcance y cubrir lo común en lo concerniente a la seguridad de todas las operaciones de la compañía.

- **Listas de chequeo del Centro para la Seguridad de Internet (CIS)**

Las listas de chequeo del CIS son una norma extensamente aceptada según el sitio web del CIS. Las listas de chequeodel CIS son basadas entre acuerdos de expertos y se aceptan ampliamente por las agencias del gobierno de los Estados Unidos("C.I.S.", 2014b).

Se le han otorgado la certificación CIS de Seguridad con las listas de chequeo a productos de las compañías sucesivas: BeyondTrust Software, Inc., Symantec, VMware, IBM, HP, Tenable Network Security, Inc., entre otras ("C.I.S.", 2014a).

Las listas de chequeo del CIS están disponibles para los SGBD siguientes: MySQLDatabase, FreeBSD, IBM DB2, Microsoft MS SQL Server, Oracle Database Server, Sybase ASE y otros ("C.I.S.", 2013b). Para otro conjunto de aplicaciones utilizadas para servidores como Apache y Microsoft Exchange.

Estas listas de chequeo son empleadas por los especialistas del DIS de la empresa ETECSA para realizar el proceso de auditoría de seguridad informática a los SGBD, por lo cual deben ser tenidas en cuenta en esta investigación.

- **Lista de chequeo de SANS**

El InstitutoSANS(SysAdmin, Audit, Red, Seguridad) se creó en1989 comola investigación cooperativay la organizaciónde la educación a más de165.000profesionales de la seguridadde todo el mundo("S.A.N.S.", 2014a).

El instituto define los Libros Blancos, fuente para la recopilación de información, resolución de problemas y el aprendizaje ("S.A.N.S.", 2014b).En estos libros se encuentran

las listas de chequeo paralas aplicaciones populares utilizadasen los distintos servicios como por ejemplo: para el firewall de Nokia, de SonicWall, para los servidores de aplicaciones web, navegadores, para los distintos gestores de bases de datos: MySQL para Linux, para Oracle Databases, SQL Server 2000, FreeBSD y otros ("S.A.N.S.", 2014b).

- **Las listas de chequeo de la NIST**

El Programa Nacional de Lista de Verificación (con las iniciales en inglés NCP), que se define por elNISTSP800-70Rev.2, es el repositorio de gobierno de los EE.UU. de listas de control de acceso público de seguridad (o referencias) que proporcionan una orientación detalladasobre la configuración a bajo nivelde laconfiguración de seguridadde los sistemasoperativos y las aplicaciones. El NCP ejecuta los pasos para ajustar su repositorio de listas de control alProtocolo de Automatización de Contenido de Seguridad (con las iniciales en inglés SCAP). El protocolosoportaestándaresde herramientasde seguridad para realizar automáticamentela configuración dela comprobaciónmediantelistas de control("N.I.S.T.", 2104).

El NCP contiene las listas de verificación para llevar a cabo la comprobación de la configuración de los sistemas de aplicación de la FDCC (siglas en inglés deFederalDesktopCoreConfiguration) y la configuración de los sistemas de aplicación de la USGCB (siglas en inglés deUnitedStatesGovernmentConfigurationBaseline),por medio del SCAP("N.I.S.T.", 2014).

Para el autor de esta investigación la NIST es quien ofrece la mayor cantidad de listas de chequeo de seguridad en cuanto a las categorías de producto y a los productos objetivos. Provee listas de chequeo compatibles con SCAP. Por lo que es la institución vista con mayor adelanto en el trabajo con el SCAP.

Estas listas de chequeo tienen como característica que posibilitan la descripción perjudicial de cada vulnerabilidad, pero a diferencia de las listas de chequeo brindada por el CIS y

SANS, no ofrece una descripción al auditor de cómo se podría verificar la existencia o no de la vulnerabilidad y el nivel de impacto que podría provocar.

Las listas de chequeo no ofrecen la posibilidad de realizar la evaluación del riesgo. Pero permite detectar las vulnerabilidades y las amenazas en las que están expuestos los SGBD, premisas importantes para determinar el nivel de riesgo de la seguridad de la información de la TI.

Las listas de chequeo tienen una fuerte dependencia de la opinión del auditor en el resultado del análisis delRSI en los SGBD(Piattini, M. G. V. y De Peso, E. N., 2001).

1.1.4 Sistemas para el análisis de las auditorías de seguridad informáticas

Existen varios sistemas informáticos que permiten la automatización de un grupo de actividades relacionadas a las auditorías informáticas y dentro de estas, la evaluación del RSI. En este sentido cabe mencionar los siguientes:

- Sistemas de automatización de metodologías de evaluación de riesgos de seguridad de la información. Un alto número de metodologías como las mencionadas, son soportadas por software. Los métodos que son soportados por sistemas informáticos tienen ciertas desventajas: el costo del método puede ser usualmente alto y el principal marco del proceso,es trazado por el software. De esta manera algunas variaciones necesarias de la metodología durante el proceso de análisis del riesgo no pueden ser alcanzadas (Carvalho, F. D. y Silva, E. M. D., 2006). Ejemplo de aplicaciones: Callio Secura.Es compatible conlosestándares ISO17799 y la ISO 27001 (BS 7799-2) ("E.N.I.S.A.", 2013). Pero es una herramienta muy limitada para el análisis de riesgo y un análisis de brecha más amplio(Calder, A. y Watkins, S. G., 2010). Entre otras están: CASIS, Citicus ONE y RiskWatch

360("E.N.I.S.A.", 2014; "S.I.G.E.A.", 2013a, 2013b; Calder, A. y Watkins, S. G., 2010; Citicus, 2014; Chou, T.-S., 2011; International, R., 2014a, 2014b).

- Sistemas de análisis automático de lasTI, en busca de posibles vulnerabilidades y fallos de seguridad. Pueden encontrar errores de administración y fallos de seguridad comunes, tales como la falta de actualizaciones del sistema, contraseñas caducadas, existencia de más de una cuenta de administrador o de una cuenta de invitado, qué directorios se comparten y con quién, posibles fallos en Internet Explorer, Microsoft Office, SQL Server, Internet Information Server, etc. Detecta los errores más comunes de configuración de seguridad y actualizaciones de seguridad que falten en los sistemas informáticos. Existen aplicaciones enfocadas a varios tipos de tecnologías como es el caso de Microsoft Baseline Security Analyzer("Softonic", 2014; "Technet", M., 2014). Otras están enfocadas solamente a los SGBD como es el caso de DB Audit, Secure Oracle Auditor y Auditor de Seguridad del SQL, Microsoft Forefont, MSAT("Microsoft", 2014, 2011; Bytes, S., 2014a, 2014b; Technologies, S., 2013), entre otras aplicaciones. Estos sistemas a pesar de ser capaces de detectar las vulnerabilidades de las TI, no llegan a evaluar el riesgo de seguridad de la información.
- Aplicaciones que automatizan las listas de chequeo. Tienen como finalidad evaluar el cumplimiento de las listas de chequeo. Esencialmente comparan las configuracionesde seguridad de los sistemas de TI según las listas de chequeo. Una de estas herramientas es el CIS-CAT, el cual trabaja según las listas de chequeo del CIS (de inglés: Center for Internet Security)("C.I.S.", 2013a) y brinda una puntuación de los parámetros cumplidos de la lista de chequeo pero no evalúa el riesgo de seguridad de la información.

Limitantes del análisis del riesgo de seguridad de la información

A pesar deldesarrollo de muchas soluciones a los problemas de la seguridad en los sistemas de información, la apreciación general es que la inseguridad es un problema que no ha sido resuelto(García, J. M., 2008).Una de las principales razones por las cuales los problemas de seguridad informática no han sido resueltos es la aparición frecuente de nuevasamenazas. Precisamente una de las debilidades de las metodologías de análisis de riesgo es que parten de una visión estática de las amenazas, así como de los controles requeridos para disminuir el riesgo.

Todo sistema de información evoluciona, debido a la integración de hardware y software con características nuevas y más atractivas para los usuarios, así como al desarrollo de nuevas funcionalidades. Estos cambios, abren la posibilidad de riesgos imprevistos y también pueden crear vulnerabilidades donde antes no existían.

1.2 Soluciones basados en el conocimiento

En un mundo donde la complejidad y los grandes volúmenes de información hacen más difícil el manejo del conocimiento, el uso de los Sistemas Basados en el Conocimiento se ha convertido en una estrategia importante para las organizaciones (Febles, O. D., Estrada, V. S., y Febles, J. P. R., 2012).

En términos generales, un Sistema Basado en Conocimiento puede ser definido como un sistema computarizado que usa conocimiento sobre un dominio para arribar a una solución de un problema de ese dominio. Esta solución es esencialmente la misma que la obtenida por una persona experimentada en el dominio del problema si se enfrenta al mismo problema (Lio, D. G., 1998).

Este tipo de sistema ayuda a la solución de problemas al recurrir a una representación simbólica del conocimiento en un determinado dominio. Se puede usar particularmente en

aquellas áreas que cuentan, como soporte básico, con el conocimiento de expertos como son la medicina, la industria, la educación, entre otras.

Los Sistemas Basados en Casos (SBC) son una de las tecnologías actuales para construir Sistema Basado en el Conocimiento para la toma de decisiones. Estos sistemas utilizan el razonamiento basado en casos (RBC) como método de solución de problemas para resolver nuevas situaciones.

Un Sistema Basado en Casos para la toma de decisiones como un Sistema de Decisión S que se define en términos de un par (U, XU Y)donde U es un conjunto finito no vacío de objetos o eventos llamados casos, mientras X e Y son dos conjuntos finitos, no vacíos de atributos o propiedades llamados conjuntos de rasgos predictores y objetivos respectivamente(Cuadrado, S. R. y otros, 2011).

$$S = (U,\ X \cup Y)$$

Cada rasgo predictor $x_i \in X$ puede ser considerado una función que mapea elementos de U en un conjunto $M_i = \{x_{i1}, x_{i2}, ..., x_{i\eta_i}\}$, el cual se denomina conjunto de valores del rasgo predictor x_i.

$$x_i : U \to M_i$$

De manera análoga cada rasgo objetivo $y_j \in Y$ puede ser considerado una función que mapea elementos de U en un conjunto $N_j = \{y_{j1}, y_{j2}, ..., y_{j\theta_j}\}$, el cual se denomina conjunto de valores del rasgo predictor y_j.

$$y_j : U \to N_j$$

En este tipo de sistema, se emplea las experiencias pasadas en forma de casos almacenados en una base de casos(BC) para apoyar la toma de decisiones en situaciones actuales similares (Zhang, J., Lu, J., y Zhang, G., 2011). El enfoque que utilizan los SBC para la

adquisición de conocimiento es una de las ventajas que se le acreditan a este tipo de sistemas; pues razonan desde episodios específicos, lo cual evita el problema de descomponer el conocimiento del dominio y generalizarlo en reglas. Otras de las ventajas de los SBC están fundamentadas en la flexibilidad para representar el conocimiento a través de los casos, la organización de la base de casos (BC) y de las estrategias de recuperación y adaptación de los casos y que el usuario puede ser capaz de agregar nuevos casos a la base de conocimiento sin la intervención experta (Sánchez, N. M., Lorenzo, M. M. G., y Valdivia, Z. Z. G., 2009).

1.2.1 Componentes de un SBC

Los componentes de un SBC son la base de conocimiento, el módulo de recuperación de casos, el de adaptación o reutilización de las soluciones, la revisión y retención(Martínez, N. S., Ferreira, G. L., García, M. M. L., y Valdivia, Z. G., 2008).

El módulo de recuperación consta de dos etapas (Gutiérrez, I. M., Bello, R. E. P., y Tellería , A. R., 2002): la etapa de acceso y la etapa de recuperación. El algoritmo de acceso a los casos debe ser rápido y eficiente. Este depende de las técnicas de indexación aplicadas y su diseño se vuelve un aspecto crítico cuando la base de casos es muy grande.

Las técnicas de indexación garantizan una recuperación eficiente de los casos, entendiéndose por eficiencia tantos los aspectos relacionados con el tiempo de recuperación como la garantía de que ningún caso relevante quede afuera de la búsqueda (Wong, L. R. P. y Mauricio, D. S., 2012).

La indexación puede ser definida de forma manual por expertos y también existen varios esfuerzos para utilizar métodos de automatización (Main, J., Dillon, T. S., y Shiu, S. C., 2001).

CAPÍTULO 1. MARCO TEÓRICO REFERENCIAL DE LA EVALUACIÓN DEL RIESGO DE SEGURIDAD DE LA INFORMACIÓN

La presente investigación se enfocó en el trabajopublicado en (Wong, L. R. P., Mauricio, D. S., y Papa, E. A., 2014), en el cual se propone un algoritmo de indexación que organiza los casos según un ordenamiento predeterminado y la condición que debe cumplirse para la clasificar es de igualdad. Aunque después de obtenida la clasificación de los casos, no tratade organizarlos.

El proceso de recuperación consiste en determinar los casos de la base más semejante a cada nuevo caso (Cuadrado, S. R. y otros, 2011). No existe una medida de semejanza única y general para cualquier dominio, sino que existen varios criterios de funciones de semejanzas(Bello, R. y Morell, C., 2000), (Morell, C. a. P., Bello, R. P., y Grau, R. Á., 2005) y (Rodríguez, Y. S., García, M. M. L., y De Baets, B., 2008).

Los algoritmos de recuperación más investigados, por el momento, son los k-vecinos más cercanos o Nearest-neigborretrieval (k-NN), árboles de decisión y sus derivados (Herrera, J. a. Q., Eduardo, J. O., Alfaro, L. C., y Tupac, Y. V., 2013).

Después de la determinación de los casos más semejantes por el algoritmo de recuperación, las soluciones contenidas en dichos casos pueden usarse directamente como solución al nuevo problema, pero comúnmente necesitan ser modificadas. Existen métodos y reglas de adaptación para realizar dicha modificación (Kolodner, J. L., 1993), (Bonzano, A., 1998) y (Mitra, R. y Basak, J., 2005).

La adaptación nula, uno de los métodos existentes se basa en no hacer adaptación, se considera una de las opciones viables porque se plantea un sistema simple en que las opciones están acotadas (García, F. G., González, P. a. C., y Sanchez, A. A., 2011).Característicaexistente en el rasgo objetivo de la BC de esta investigación por lo que es el tipo de adaptación escogida.

Una vez realizada la adaptación, es importante evaluar el caso dado como respuesta(Lawanna, A. y Daengdej, J., 2010). La misma puede ser dada por un experto o semiautomatizada(Navarro, M. I., 2011). Una vez que fue evaluado se procede a la reparación en caso que sea necesario. Como la adaptación escogida es la nula, la reparación es simple, es cambiar el valor proporcionado en caso de que sea incorrecto. El auditor debe tener la última decisión en el proceso de evaluación, por eso se deja en manos de este, la tarea de revisión.

En el componenteretención, se encuentra el aprendizaje y el mantenimiento de la BC. La descripción del problema y su solución puede ser entoncesretenida como un nuevo caso y el sistema aprende a resolver unnuevo problema.

Para dotar al SBC de una capacidad inteligente es muy importante incorporar la etapa de aprendizaje (Navarro, M. I., 2011). La reutilización tiene un papel fundamental en este proceso. Se ha llegado a argumentar que la reutilización es el paso más importante del RBC, debido a la incorporación inteligente,sino sería un mero proceso de reconocimiento de patrones (Lozano, L. y Fernández, J., 2006).

Existen varios algoritmos para el aprendizaje, donde se utiliza el recuerdo para actualizar la experiencia de un sistema que usa un modelo dinámico de aprendizaje reforzado(Salamó, M. y Golobardes, E., 2004). El aprendizaje reforzado es un acercamiento a aprender por el ensayo y error para lograr una meta(Harmon, M., 1996).

En la fase de aprendizaje está presente el mantenimiento (Haouchine, M.-K., Chebel-Morello, B., y Zerhouni, N., 2008). El mantenimiento del sistema controla el crecimiento de la base de casos, puede influir en el rendimiento del mismo (Navarro, M. I., 2011). En el mantenimiento son importantes las acciones de eliminar, añadir y revisar los casos. Por lo que los algoritmos desarrollados para el mantenimiento están imbricados con el

aprendizaje. De esta forma se pueden ver algoritmos de aprendizaje manifestados como de mantenimiento con clasificaciones.

Se clasifican los métodos de mantenimiento de una base de casos en dos ramas: por una parte los que siguen una política de optimización, los cuales utilizan algoritmos sofisticados para agregar o eliminar los casos para perfeccionar la base de casos. Por otro lado están los que siguen una política de particionamiento, las cuales permiten construir una elaborada estructura de la base de casos y posibilitan un mantenimiento continuo(Smiti, A. y Elouedi, Z., 2014).El rendimiento depende críticamente de la exactitud y los casos guardados en labase de caso.

En la política de optimización se pueden utilizar los casosselectivamentede una BCguiadapor esta clasificaciónde los casos, para el proceso del mantenimiento.

El modelo WCOID (Smiti, A. y Elouedi, Z., 2014)surge como extensión del modelo COID (Smiti, A. y Elouedi, Z., 2010; Smiti, A. y Elouedi, Z., 2014). Este es un ejemplo de algoritmo con política de particionamiento, muy utilizado para escenarios de la existencia o necesidad de clúster, pero para esta investigación no se presenta esta necesidad. Sin embargo, el autor considera útil la política de particionamiento para una más rápida recuperación de los casos y por tanto el algoritmo de aprendizaje debe tenerla en cuenta para la creación de la estructura de la BC.

Modelo de la base de casos

La BC contiene las experiencias, ejemplos o casos a partir de los cuales el sistema hace sus inferencias. Estas bases pueden ser generadas por entrevistas a expertos humanos o por un procedimiento automático o semiautomático que construye los casos desde datos existentes, registrados por ejemplo, en una base de datos (Sánchez, N. M. y otros, 2009).

Un razonador basado en casos depende de la estructura y el contenido de la base de conocimiento.

Para resolver el problema de organizar una BC, un enfoque ha sido almacenar los casos de forma secuencial o plana y analizarlos todos para resolver el nuevo problema.

La estructuración de la BC como memoria planaes más sencilla, pero la búsqueda de casos hace lento el proceso de recuperación(Juárez, J. y J., P., 2005).

Un método alternativo es una memoria jerárquica. En este caso, la estructura de almacenamiento es un grafo, donde cada nodo almacena ciertos atributos compartidos o no por varios casos. Normalmente, se suele utilizar estructuras de árbol, donde los nodos interiores son atributos comunes y las hojas contienen los atributos diferenciados. Así, el camino desde el nodo raíz a la hoja establece todos los atributos del caso, tanto los comunes como los diferentes, que distinguen un caso de otro(Juárez, J. y J., P., 2005). A la hora de establecer la estructura arbórea, se selecciona como nodo raíz aquel atributo más discriminante.El autor de esta investigación considera que esta estructura puede aportar a la indexación de los casos para el caso de esta investigación, teniendo en cuenta el resultado del trabajo publicado porMartínez, Ferreira, García y Valdivia(2008) y además permite aplicar una política de particionamiento.

Usar una jerarquía abstracta donde cada nodo es una abstracción de los casos representados por sus hijos es otra de las propuestas de modelos de BC (Kolodner, J. L., 1993). Estas jerarquías se conocen como redes de discriminación donde los nodos representan regiones de superposición de casos. Sin embargo, este enfoque requiere más memoria para almacenar la red y los procedimientos para agregar nuevos casos son muy caros; la jerarquía de abstracción necesita ser reestructurada cada vez que se incorpora un nuevo caso.

Otro método, son los modelos basados en ejemplares que no necesariamente requieren todos los rasgos o todos los casos por adelantado. Sin embargo, buscar un criterio para determinar cual es un buen ejemplar, no resulta una tarea trivial.

Descripciones de los enfoques mencionados anteriormente y referencia a modelos donde son utilizados pueden encontrarse en Gutiérrez y Bello(2003) y en López (2005).

1.2.2 La lógica difusa

La lógica difusa fue introducida para la manipular de ambigüedades, de valores inciertos e imprecisos existentes en la vida real(Zhang, J. y otros, 2011). El concepto que introdujo Zadeh al cual llamó "Razonamiento Aproximado", con el cual demostró que los estados lógicos vagos permiten la formación de los algoritmos que pueden utilizar datos vagos para derivar inferencias imprecisasy asume que su enfoque sería beneficioso, sobre todo, en el estudio de los sistemas humanísticos complejos(Sarkar, A., Sahoo, G., y Sahoo, U. C., 2012).

En la actualidad existen varios trabajos que relacionan el RBC con la lógica difusa como se mencionan en las tesis doctorales de Morell(2005) y Rodríguez (2008), aplicadas a atributos o rasgos con valores imprecisos y valores lingüísticos.

El concepto fundamental de los sistemas borrosos es la noción de grado de pertenencia (Dubois, D., Prade, H., y Klement, E. P., 1999), generaliza la idea de que un elemento pertenezca o no a un conjunto, a una pertenencia gradual mediante la fusificación del predicado "pertenece" (es elemento de). Para determinar el grado de pertenencia, se expresan de varias maneras tales como: en funciones Trapezoidales, Gaussianas, Campanas y Triangulares.

Un sistema de inferencia borroso (SIB) es un sistema computacional basado en los conceptos de la teoríade conjuntos borrosos, reglas If-Then borrosas y razonamiento

borroso(Bello, R. P., 1998).Existen tres tipos fundamentales de sistemas de inferencia borrosos: el modelo de Mamdami, Sugeno y Tsukamoto que se diferencian en la forma del consecuente de sus reglas borrosas (Piñero, P. Y. P. y García, M. M. L., 2005)

Las entradas y salidas pueden ser valores duros o borrosos. Si la salidaes borrosa y se necesita el valor duro, se emplea un método de defusificación que determina el valor duro que mejor representa un conjunto borroso. El proceso de computarun escalar desde una conclusión difusaque es una variable lingüística,cuyos valores le han sidoasignados grados depertenencia, es llamado comodefusificación(Siler, W. y Buckley, J. J., 2005). En ese libro se mencionan varios métodos de defusificación, incluido el del centroide como el más utilizado, pero existen otros métodos como: el bisector del área; el medio de máximos: es el promedio de los valores del universo de discursodonde se alcanza el máximo grado de membresía; el más pequeño de máximos: es el valor menor del universo de discurso con el cual se alcanza el máximo grado de membresía y el más grande de máximos: es el valor mayor del universo de discurso con el cual se alcanza el máximo grado de membresía.

Es debido a las facilidades de esta técnica que se han desarrollado varias investigaciones con resultado donde se vincula la evaluación del riesgo en variadas áreas del conocimiento con la lógica difusa donde pueden hallarse nuevas propuestas de funciones de similitud para números difusos (Schmucke, K. J., 1984), (Patra, K. y Mondal, S. K., 2015), (Abbasianjahromi, H. y Rajaie, H., 2013),(Wang, Y.-M. y Elhag, T. M., 2006) y (Xu, Z., Shang, S., Qian, W., y Shu, W., 2010). Debido a ello es que el autor de la presente investigación incorpora decida valorar la posibilidad de introducir la lógica difusa para en el proceso de evaluación de RSI en los SGBD.

1.2.3 Sistemas inteligentes para el análisis de riesgo de seguridad de la información

BDSS: Sistema de Soporte de Decisión Bayesiana (con las siglas en inglés) y el **ALRAM**Automatizada Metodología (con las siglas en inglés) aplican algoritmos de evaluación del riesgo de la industria nuclear para la seguridad de la información, y remplazan el algoritmo FIPSPUB-65. BDSS es un sistema experto que proporciona al usuario una base de conocimiento amplia que dirige las vulnerabilidades y los resguardos, así como los factores de exposición de las amenazas y la frecuencia de los datos. Todos son totalmente mapeados y cruzados con algoritmos de modelación del riesgo e interfaces y presentación de lenguaje natural. ALRAM, sin embargo, requiere un experto para construir y mapear la base de conocimiento y entonces conducir a una evaluación del riesgo personalizada (Peltier, T. R., 2001).

Estos sistemas aunque están enfocados al riesgo de seguridad en el área de la industria nuclear y no a los sistemas informáticos, proponen a través de técnicas de inteligencia artificial una forma de evaluar, por lo que exponenun camino para que puedan ser utilizadas las mismas en el área de la seguridad informática. Conociendo que el valor del riesgo es un número probabilístico, una solución a través de un modelo bayesiano se puede presentar idóneo. El problema en este enfoque es prevenir los suficientes datos a priori que permitan realizar la estimación probabilística del fenómeno. Si se encuentra una auditoría donde las características la convierten en única o nueva, se podría entrar en dificultades para utilizar el modelo bayesiano.

MIDS: Es un Sistema de Detección de Intruso (con las siglas en inglés). Tiene como principal ventaja una proporción baja de falsos positivos y es capaz de realizar acciones preventivas y correctivas. El sistema puede ser dividido en varias categorías, una de ellas es como sistema experto. En esta categoría el sistema tiene un conjunto de reglas que

describen los ataques empleados. Los eventos de auditorías se traducen como hechos, llevan su significación semántica en el sistema experto y el motor de inferencia traza conclusiones al usar las reglas y hechos (Chou, T.-S., 2011).

La baja proporción de falsos positivos demuestra la efectividad de utilizar las reglas y hechos en el área de la seguridad informática, en la detección de intrusos, por lo que a pesar de que este sistema no está pensado para la evaluación del riesgo de seguridad de la información, sí demuestra la efectividad de estas técnicas en este escenario, al proporcionar propuestas de soluciones a los problemas planteados en este trabajo.

@RISK: este sistema realiza análisis de riesgo al utilizar la simulación para mostrar múltiples resultados posibles en un modelo de hoja de cálculo y le indica qué probabilidad existe de que se produzca. Computa y controla matemáticamente, un número de escenarios futuros e indica las probabilidades y riesgos asociados con cada uno. Esto quiere decir que permite decidir qué riesgos tomar y cuáles evitar y tomar la mejor decisión en situaciones de incertidumbre(Palisade, 2014).El sistema @RISK utiliza la simulación de Monte Carlo para el cálculo del riesgo. El mismo utiliza algoritmos genéticos para determinar la mejor asignación de recursos y la distribución óptima de activos (Palisade, 2014).

El método de Monte Carlo está basado en números generados al azar. Los escenarios que construye están basados en estos números. La probabilidad de ocurrencia de un ataque a las vulnerabilidades a un sistema gestor de base de datos no es obra del azar, sino de un estudio del atacante. Por lo que, si las vulnerabilidades conocidas están fuertemente protegidas o eliminadas, puede repercutir en que la frecuencia de ocurrencia de un ataque puede ser cero o muy baja, por la protección existente. Ante esta situación puedeconsiderarse que el riesgo es casi nulo, mientras que los números aleatorios pueden estar proponiendo otros valores.

1.2.4 Otras soluciones basados en técnicas de inteligencia artificial

En esta investigación el autor ha identificado varias soluciones que utilizan técnicas de inteligencia artificial para la evaluación del riesgo en otras áreas del conocimiento.

La propuesta publicada porZhang, J., Lu, y Zhang, G(2011)tiene un método de estimación que emplea la lógica difusa y el RBC. A través de estas técnicas se logra generar advertencias tempranas que apoyen a quienes toman decisiones, para identificar si están bajo vulnerabilidades y para instrumentar las estrategias en los sistemas de advertencia pertinentes en la influenza aviar.La propuesta propone en esencia calcular el riesgocon cada una de las técnicas mencionadas y seleccionar el resultado mayor entre los dos.

El autorde esta investigación considera que esta solución permite utilizar las experiencias pasadas y manejar las ambigüedades de los valores lingüísticos de la salida de esta propuesta, lo cual resulta de utilidad para la evaluación del RSI. Como limitación, se aprecia que el modelo utiliza para determinar el riesgo, valores cuantitativos, entradas que no se ajustana las condicionesdela presente investigación.

Se considera una debilidad de esta propuesta calcular dos veces el riesgo para posteriormente elegir un resultado. Al utilizar las dos técnicas de IA por separado, aumenta la complejidad computacional. Tiene a favor la demostración de la viabilidad de utilizar ambas técnicas de la IA para obtener la evaluación del riesgo.

Martell y Zulueta (2014)publicaron un modelo para desarrollar el proceso de análisis de riesgos en líneas de productos de software. La misma incluye una solución para la toma de decisiones multicriterio y dinámica, para evaluar los riesgos con influencia sobre uno o varios activos. En esta propuesta, el análisis de un activo, puede influir sobre otro. Incluye las evaluaciones históricas de los riesgos y el concepto de facilidad de detección de un

riesgo. De esta forma se utilizan las evaluaciones de los riesgos de las auditorías de seguridad informáticas anteriores, en el proceso de evaluación presente.

Para el caso que trata el modelo, el impacto del riesgo, proviene del cálculo de los activos que pueden sufrir el riesgo. El impacto de este último se mide comparándolo con el impacto de otros activos. Se entiende que el impacto de cada activo guarda una relación con los otros, pero para el caso del impacto que puede tener cierta vulnerabilidad en un SGBD, el mismo no estáligado a otras vulnerabilidades. Aunque puede existir cierta relación entre las vulnerabilidades que están enfocadas hacia un mismo tipo de debilidad en el sistema. Por ejemplo: para los permisos de la base de datos, se pueden tener varias vulnerabilidades en los permisos de acceso a la base de datos, en los privilegios de los usuarios de los administradores y otras más. Pero estas no guardan ninguna relación con las vulnerabilidades identificadas en las configuraciones SSL. Por lo que el modo de calcular el impacto en el modelo analizado no se ajusta a las condiciones de esta investigación.

En este modelo se exponen dos variantes o métodos para la obtención del riesgo. La fórmula de exposición del riesgo propuesta en el primer método es a través de la suma ponderada.Este método no se ajusta para el cálculo del riesgo de seguridad de la información según los métodos cuantitativos de análisis del riesgo. En el modelo se plantea una sumatoria de la variable facilidad de detección del riesgo de forma pesada, en conjunto a la sumatoria de la variable probabilidad, también pesada, con el impacto, igualmente acompañada de un peso. Esta fórmula es desigual a varios métodos cuantitativos de evaluación de riesgo de seguridad de la información, donde exponen que el riesgo está constituido por el impacto y la probabilidad de ocurrencia de la vulnerabilidad. La facilidad con que se detecta un riesgo no es una variable que forme parte del riesgo. Por lo

que la fórmula de exposición de riesgo presentada por el modelo no es prudente su uso para el problema planteado en el área de la evaluación del riesgo de seguridad de la información.

El cálculo de la exposición del riesgo a través del segundo método del modelo, a través del análisis dinámico, utiliza la estimación del riesgo en otras iteraciones en el mismo proyecto. Es en esta variante donde se utilizan evaluaciones anteriores. El modelo debe calcular el riesgo varias veces, según se vayan ejecutado las acciones de mitigación y/o contingencia. Sin embargo, no se plantea la posibilidad de la utilización de los riesgos calculados en otros proyectos similares, como medida de ajuste de exposición al riesgo del actual proyecto. Por lo que la utilización de riesgos históricos se reduce a los mismos riesgos calculados anteriormente durante la existencia o vida del proyecto.

Por lo planteado anteriormente se considera que no debe ser utilizado para la evaluación del riesgo de la seguridad de la información en los SGBD.

En eltrabajode Hernández, Yelandy y Cuza (2013)se propone el uso de los modelos causales para realizar el análisis de los riesgos basado en la información obtenida del proyecto o empresa. Dígase riesgos y la influencia de los mismos sobre los costos, tiempo de desarrollo y alcance del proyecto, para así detectar las afectaciones de un sin números de tareas que tributan a su desarrollo.

Específicamente en este modelo se hace uso de los mapas cognitivos difusos (MCD), que se construyen a partir de una base de conocimiento basada en los riesgos. Es decir, que los MCD parten de una base de riesgos previamente calculados. A partir de este, se construye un grafo ponderado con las relaciones entre los riesgos, que serían los nodos y el peso de los arcos, es el nivel de relación entre los nodos. De esa forma se logra obtener la

priorización de los riesgos y determinar el riesgo más perjudicial. El modelo proporciona como resultado el peor caso que puede suceder, el más óptimo y el promedio.

Para el problema planteado en este trabajo, este modelo no aporta beneficios ya que no tiene como objetivo la evaluación del riesgo, sino parte de esta, con el fin de cumplir su propia meta.

El trabajo publicado por Abbasianjahromiy Rajaie(2013)permite tomar decisiones a partir de factores de riesgo de varios proyectos de la construcción utilizando un modelo delRBC difusocon factores de riesgos cualitativos como rasgos predictores, paradecidir cuál proyecto ejecutar, utilizando una nueva función de semejanza basado en números difusos trapezoidales. Esta solución es considerada una propuesta viable al ajustarse a las condiciones de esta investigación, por lo que se tiene en cuenta en el proceso de evaluación del RSI.

Los modelos publicados porVicente, Mateos y Jiménez (2013) y Patra y Mondal(2015) proponen soluciones que tienen como rasgos predictores, el riesgo con valores cualitativos, en las cuales se proponen nuevas funciones de semejanza con mejores resultados a los existentes anteriormente en donde se utilizan números difusos trapezoidales. El autor decide utilizar la función de semejanzapropuesto porVicente, Mateos y Jiménez (2013) por la superioridad, según las comparaciones publicadasen este mismo trabajo con las siguientes funciones de similitud que manejan números difusos trapezoidales:(Chen, S.-J. y Chen, S.-M., 2003), (Nayagam, V. L. G. y Sivaraman, G., 2012), (Zhu, L.-S. y Xu, R.-N., 2012), (Wei, S.-H. y Chen, S.-M., 2009), (Sridevi, B. y Nadarajan, R., 2009), (Xu, Z. y otros, 2010), (Hsieh, C. H. y Chen, S. H., 1999) y (Hejazi, S. R., Doostparast, A., y Hosseini, S. M., 2011).Aunque en Patra y Mondal(2015) se publica la superioridad de supropuesta ante estas mismas funciones, la publicada en Vicente, Mateos y Jiménez

(2013), obtienen mejores resultados en los mismospares de números difusos trapezoidales con los cuales se realizaron la comparación.

1.4 Conclusiones parciales

La revisión de la literatura y el desarrollo consecuente del marco teórico, arrojó que actualmente existen variosmecanismos o víasrelacionados con la evaluación delRSIen los SGBD pero que no resuelven el problema plantado en esta investigaciónrespecto al aumento de la exactitud y la disminución del tiempo de respuesta.

Es necesaria una solución que permita realizar una evaluación del RSI para los SGBD que permita manejar la ambigüedad de los términos lingüísticos ycontribuya a determinar con mayor precisión el resultado del riesgo,al aprovechar la experiencia de los expertos y reduzca el tiempo de respuesta en el diagnóstico. Se determinó utilizar varias soluciones encontradas en las cuales utilizan el RBC y la lógica difusa para hallar el riesgo en otras áreas del conocimiento. La decisión se debe a que tiene en cuenta la experiencia de las personas, las decisiones tomadas en situaciones similares, el empleo de términos lingüísticos como entrada y en la salida de respuestay la viabilidad mostrada para determinar el riesgo en otras áreas del conocimiento como se comprobó en los trabajos analizados.

CAPÍTULO 2

MODELO PARA LA EVALUACIÓN DEL RIESGO DE SEGURIDAD DE LA INFORMACIÓN EN LOS SISTEMAS GESTORES DE BASES DE DATOS

CAPÍTULO 2. MODELO PARA LA EVALUACIÓNDEL RIESGO DE SEGURIDAD DE LA INFORMACIÓN

En este capítulo se analizan los resultados del diagnóstico aplicado a una muestra representativa de especialistas en la auditoría de seguridad de la información de la empresa ETECSA en Cuba. Se exponen los aspectos más significativos del modelo. Se describen las características esenciales que identifican la propuesta y se establece la forma de realizar la evaluación del RSI en los SGBD.

En el estudio efectuado colaboraron auditores con experiencia en las auditorías de seguridad de la información en los sistemas de cómputo pertenecientes a la empresa de telecomunicaciones ETECSA.

2.1 Diagnóstico sobre la evaluación del riesgo de seguridad de la información

Un diagnóstico es el proceso mediante el cual se lleva a cabo un análisis para recopilar información que ayude a determinar la situación actual de la organización y detectar sus áreas de mejoramiento. Mediante un diagnóstico se trata de focalizar y evaluar un conjunto de variables que juegan un importante papel en la comprensión, predicción y control del comportamiento de un fenómeno determinado (Kelly, A. E., Lesh, R. A., y Baek, J. Y., 2014).

En la actualidad la evaluación del RSI en los SGBD constituye una tarea engorrosa, compleja y en lo fundamental basa sus resultados en las experiencias y opiniones de los auditores. Existen herramientas que apoyan el proceso de auditoría pero no llegan a realizar una evaluación del RSI a los SGBD, por ello, el objetivo principal del diagnóstico es caracterizar el proceso de evaluación del riesgo en las auditorías de seguridad informática en los SGBD, evaluar cómo se lleva a cabo el proceso de auditoría en ETECSA e identificar las principales debilidades y/o limitaciones de este proceso.

2.1.1 Aplicación de la encuesta

Como parte del diagnóstico se aplicó una encuesta (anexo 1) a 19de 33 especialistas que pertenecen al Departamento de Seguridad Informática de ETECSA en todo el territorio nacional. Las características de los expertos encuestados se encuentran en el anexo 2, lo que representa aproximadamente el 57 % del total.El objetivo de la encuesta fue caracterizar el proceso de evaluación del riesgo en las auditorías de seguridad informática a las TI de la empresa. Los resultados principales de la encuesta se muestran a continuación:

- Aproximadamente el 68 % de los encuestados afirman que su formación es empírica.
- El 42 % de los encuestados afirmó que no utilizan una metodología para la evaluación del riesgo de seguridad de la información por lo que la experiencia y la opinión de cada experto incide fuertemente en el resultado de la auditoría a los SGBD.
- El 68,4 % de los encuestados considera que puede existir diferencias en la evaluación del riesgo de seguridad de la información en dependencia de los especialistas que participan, al punto de llegar a cambiar el resultado de la evaluación.
- El 68 % afirmó que es demorado el proceso para obtener el resultado de la evaluación del riesgo a las TI a partir del inicio de la auditoría.
- El 100 % de los auditores afirma que sería más efectiva la evaluación del riesgo de seguridad de la información si se pudiera disponer, en el momento de la auditoría, de información automatizada sobre auditorías pasadas que pudieran servir de referencia y así apoyar la toma de decisiones.

- El 100 % considera importante la fundamentación matemáticas en la toma de decisiones para determinar el riesgo de seguridad de la información.
- El 100% de los auditores expresaron que los resultados de la auditoría se proporcionan en términos cualitativos, lo que resulta un poco ambiguo.
- El 100 % indicó que realizan las auditorías utilizando las listas de chequeo de seguridad, por lo que el resultado final del diagnóstico tiene reflejado la experiencia del auditor.

2.1.2 Entrevistas

Como parte de la presente investigación se entrevistaron a 6 especialistas con experiencia en auditorías de seguridad informática (anexo 1A). Los principales resultados obtenidos de esas entrevistas se refieren a continuación:

- La experiencia y puntos de vista del auditor tiene alta influencia en el resultado final en la evaluación del RSI.
- Existen diferentes niveles de experticia entre los auditores. La subjetividad incide negativamente en los niveles de exactitud
- Las listas de chequeo tienen una fuerte dependencia de la opinión del auditor en el resultado del análisis del RSI en los SGBD.
- La evaluación del RSI en los SGBD se expresan en los términos: Alto, Medio o Bajo, por lo que para cada auditor, constituye una medida ambigua, sin límites precisos.
- En ocasiones los criterios emitidos en la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos no son adecuados.
- Resulta conveniente utilizar las auditorías pasadas para apoyar las evaluaciones del riesgo de seguridad de la información en los sistemas gestores de bases de datos.

2.1.3 Análisis documental

Los principales documentos analizados fueron referidos en el capítulo 1. Un resumen de ello se refleja a continuación:

- El análisis de las metodologías de análisis del riesgo de seguridad de información en el epígrafe 1.1. determinó que son generales, por lo que es difícil evaluar el RSI en entidades con diferentes características de seguridad.

- Se examinó el protocolo SCAP y se concluye que no es capaz de evaluar el riesgo de seguridad de la información aunque se puede chequear las configuraciones de seguridad y realizar reporte de las mismas.

- La revisión del estándar ISO 27005 concluye que la misma no proporciona una metodología de análisis delriesgo, sino que describe a través de su conjunto de cláusulas el proceso recomendado de análisis y las fases que lo conforman.

- Se analizaron las listas de chequeo del CIS, de la SANS y de la NIST. Las listas de chequeo no ofrecen la posibilidad de realizar la evaluación del riesgo. Pero permite detectar las vulnerabilidades y las amenazas en las que están expuestos los SGBD. Las listas de chequeo tienen una fuerte dependencia de la opinión del auditor en el resultado del análisis del RSI en los SGBD (Piattini, M. G. V. y De Peso, E. N., 2001).
- Se exploraron varias herramientas informáticas como las mencionadas en el epígrafe 1.1.4, pero estas no llegan a evaluar el riesgo de seguridad de la información en los SGBD.

2.2 Modelo para la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos

La palabra modelo proviene del latín *modulus* que significa medida, ritmo, magnitud y está relacionada con la palabra *modus* que significa copia, imagen(Del Valle, A. L., 2007).

La construcción conceptual del modelo propuesto en la presente investigación constituye una combinación de teoría y práctica. A continuación se expone el modelo teórico propuesto teniendo en cuenta su objetivo, principios y premisas.

El modelo describe y representa sus componentes y las interrelaciones existentes entre ellos, que contribuyen a su vez a la evaluación del RSI con impacto en el tiempo de respuesta y en la exactitud de los resultados.

El modelo integra los principales fundamentos teóricos abordados en el capítulo 1 de la presente investigación, además de los elementos obtenidos a partir del diagnóstico realizado en ETECSA. El modelo propuesto se apoya en los conocimientos y experiencias pasadas para la toma de decisiones reflejadas en auditorías anteriores.

2.2.1 Principios y premisas del modelo

El modelo se regirá por los siguientes principios:

- La actualización permanente mediante la incorporación de los nuevos casos presentados.
- La flexibilidad para ajustar las variables del riesgo según lo determinen los auditores expertos.
- La estandarización del procedimiento de auditoría a SGBD para la evaluación cualitativa del riesgo.
- La interoperabilidad entre los componentes que conforman el modelo.

Las premisas del modelo propuesto son:

- Disponer de la lista de chequeo de seguridad del CIS para su funcionamiento como entrada.
- Identificar el tipo y la versión del gestor de base de datos con propósito de auditar.
- Los auditores deben revisar los valores cualitativos del riesgo local de los parámetros de la lista de chequeo para corregir alguna imprecisión.

2.2.2 Componentes del modelo

El modelo está integrado por componentes relacionados entre sí como se muestra en la figura 2.1.

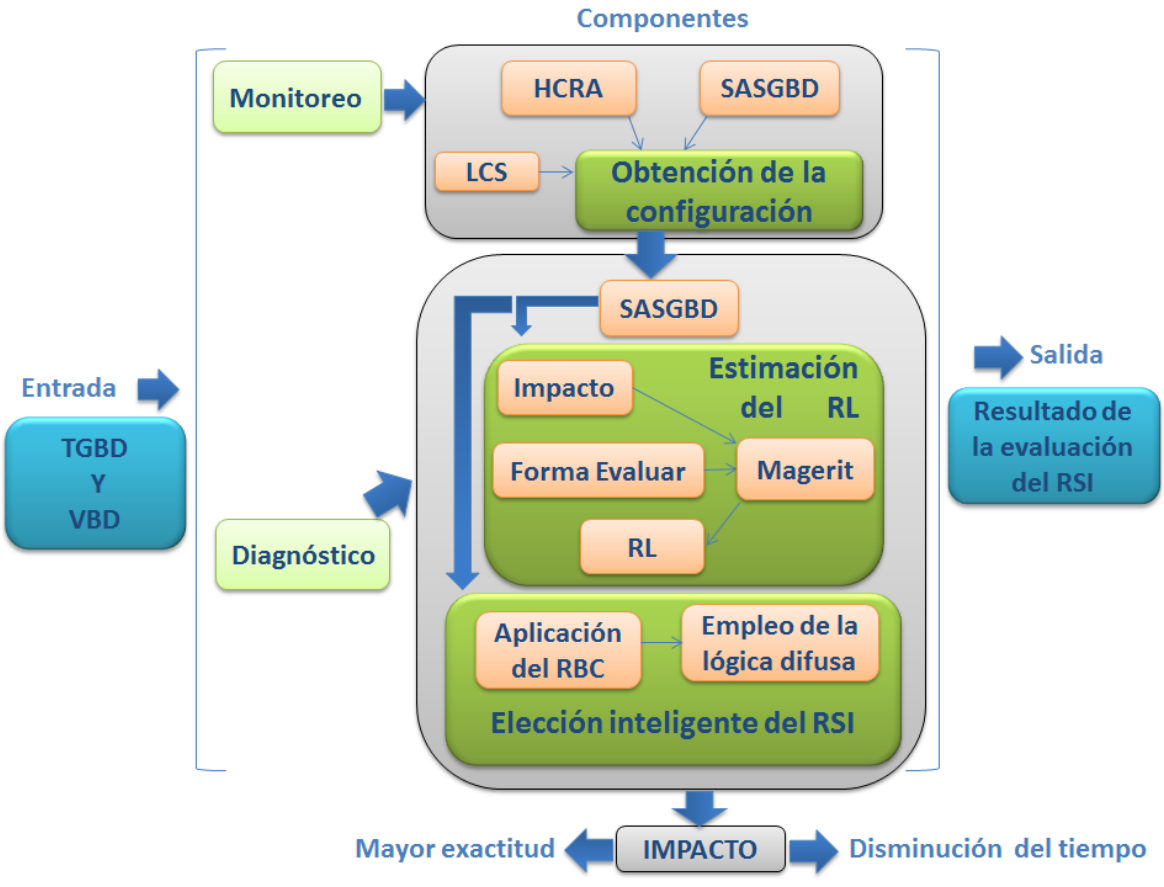


Figura 2.1. Representación gráfica del modelo.

El modelo está formado por componentes relacionados entre sí como se muestra en la figura 2.1 y agrupados por las fases: Monitoreo y Diagnóstico.

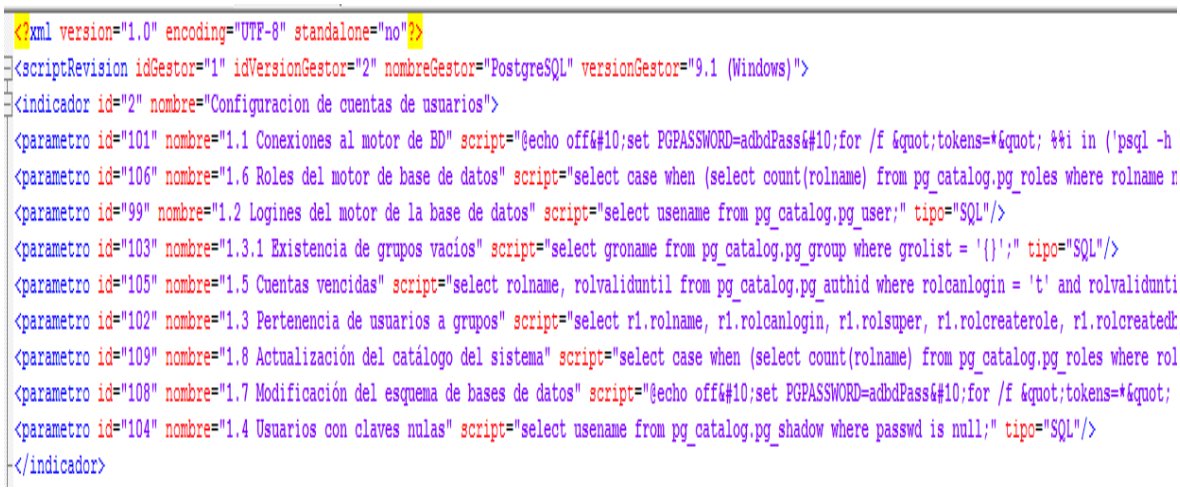
Los componentes están en diferentes fases porque están en diferentes momentos de la auditoría con diferentes requisitos y lugar de trabajo. En la fase Monitoreo, el auditor con la presencia del administrador del SGBD, sustrae las configuraciones de seguridad desde el local donde se encuentra. Para sustraer las configuraciones de seguridad, se utiliza la lista de chequeo de seguridad del CIS correspondiente al TGBD y a la VBD. Se requiere la presencia del administrador del SGBD para que otorgue las credenciales necesarias al auditor para que pueda realizar esta acción y además verifique que durante el periodo de monitoreo, se compruebe que el auditor no realizó una acción que pueda provocar problemas o fallos al SGBD y que los comandos, consultas ejecutadas y alguna otra herramienta puedan ser revisadas por este administrador.

La fase Diagnóstico, es un paso posterior donde no es necesaria la presencia del administrador del SGBD. Sino que se realiza preferencialmente en el local de trabajo del auditor donde puede ser auxiliado por otros auditores y donde son convocadas reuniones de trabajo para analizar la auditoría presente y realizar la toma de decisión con respecto al resultado de la auditoría a presentar.

a) Componente: Obtención de la configuración

Para el funcionamiento de este componente es necesaria la entrada de la versión de la base de datos (VBD) y el tipo de gestor debase de datos (TGBD).

El componente contiene la Herramienta Colaborativa para la Realización de Auditorías (HCRA), la cual está destinada a apoyar las auditorías que se realizan a través del SASGBD (Sistema de Auditoría para los Sistemas Gestores de Bases de Datos). El HCRA se enfoca en obtener las configuraciones de seguridad del servidor auditado. Las configuraciones están organizadas a través de los parámetros exportadosdel SASGBD por un archivo con extensión XML, como el fragmento mostrado en la figura 2.2. Los parámetros son los existentes en las listas de chequeo del CIS (Ver anexo 3).



```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<scriptRevision idGestor="1" idVersionGestor="2" nombreGestor="PostgreSQL" versionGestor="9.1 (Windows)">
<indicador id="2" nombre="Configuracion de cuentas de usuarios">
<parametro id="101" nombre="1.1 Conexiones al motor de BD" script="@echo off&#10;set PGPASSWORD=adbdPass&#10;for /f &quot;tokens=*&quot; %%i in ('psql -h
<parametro id="106" nombre="1.6 Roles del motor de base de datos" script="select case when (select count(rolname) from pg_catalog.pg_roles where rolname r
<parametro id="99" nombre="1.2 Logines del motor de la base de datos" script="select usename from pg_catalog.pg_user;" tipo="SQL"/>
<parametro id="103" nombre="1.3.1 Existencia de grupos vacíos" script="select groname from pg_catalog.pg_group where grolist = '{}';" tipo="SQL"/>
<parametro id="105" nombre="1.5 Cuentas vencidas" script="select rolname, rolvaliduntil from pg_catalog.pg_authid where rolcanlogin = 't' and rolvalidunti
<parametro id="102" nombre="1.3 Pertenencia de usuarios a grupos" script="select r1.rolname, r1.rolcanlogin, r1.rolsuper, r1.rolcreaterole, r1.rolcreatedb
<parametro id="109" nombre="1.8 Actualización del catálogo del sistema" script="select case when (select count(rolname) from pg_catalog.pg_roles where rol
<parametro id="108" nombre="1.7 Modificación del esquema de bases de datos" script="@echo off&#10;set PGPASSWORD=adbdPass&#10;for /f &quot;tokens=*&quot;
<parametro id="104" nombre="1.4 Usuarios con claves nulas" script="select usename from pg_catalog.pg_shadow where passwd is null;" tipo="SQL"/>
</indicador>
```


Figura 2.2. Fragmento de parámetros exportadosdel SASGBD a un XML.

El HCRA tiene la capacidad de cargar el archivo XML y mostrar las consultas SQL y los comandos a ejecutar que se utilizan para encuestar las configuraciones de seguridad de la base de datos como se muestra en la figura 2.3.

De este modo el administrador del SGBD conoce las acciones del monitoreo que se van a realizar a la base de datos, lo que brinda una mayor confianza con relación al trabajo de los auditores o de las acciones sobre el servidor.

Para obtener los datos deseados,es necesario que el administrador del SGBD entregue las credenciales necesarias al HCRA para la conexión al servidor para que pueda cumplirse la fase de monitoreo del modelo.

La solución informática HCRA es capaz de revisar los siguientes SGBD: PostgreSQL, MySQL, SQL Server y Oracle, que son los principales gestores auditadospor los especialistas en ETECSA.

La herramienta sustituye el trabajo de varias herramientas informáticas para la monitoreo de las configuraciones de seguridad del SGBD. Además, se unifica en un solo formato los resultados obtenidos en la ejecución de variadas instrucciones en SQL y comandos del sistema operativo.

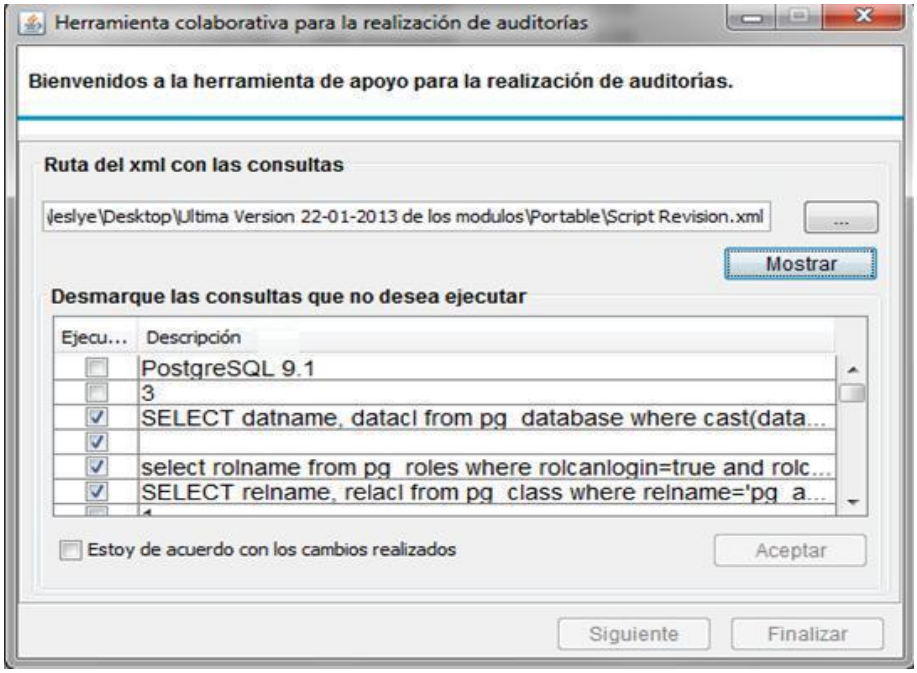


Figura 2.3. Interfaz de la aplicación HCRA.

El archivo exportado por el HCRA, se convierte en la entrada del siguiente componente del modelo propuesto en esta investigación. Con los datos de configuración introducidos, se crea una matriz de diagnóstico para cada servidor monitoreado a través de la aplicación SASGBD. La figura 2.4 muestra el proceso que realiza el componente descrito.

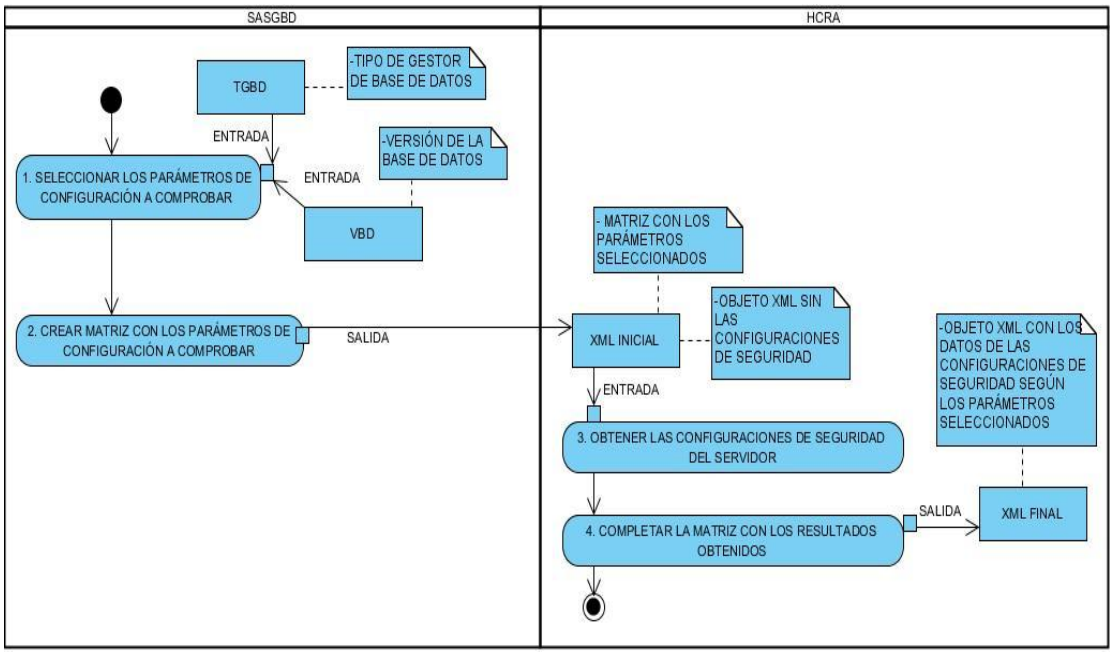


Figura 2.4. Esquema del componente Obtención de la configuración.

b) Componente: Estimacióndel riesgo local

Los especialistasencuestados en el diagnóstico estiman el riesgolocal (RL) para cada parámetro de la lista de chequeo de seguridad, solamente con valores lingüísticos: Alto, Medioo Bajo. Por eso, en ese componente para determinar el RLse utiliza una modificaciónde la técnica análisis de tabla propuesta en la metodología Magerit 3.0(M.H.A.P., 2012a), la cual se asemeja al trabajo de los especialistas.En la tabla 2.1 se muestra cómo se encuentra diseñada la modificación hecha a esta técnica,a las exigencias del proceso de auditorías señaladas por los especialistas. Se sustituye la variable probabilidad de ocurrencia de un ataque a cierto parámetro por la evaluación del parámetro (EP) dado por los especialistas.Solo se utilizan,dentro de la gama de valores lingüísticos existentes en la metodología Magerit, las utilizadas por los especialistas en las tablas 2.1 y 2.2.

Tabla 2.1. Análisis mediante tablas basado en Magerit.

RL		**Evaluación del parámetro (EP)**	
		Bien	**Mal**
Impacto	**A**	B	A
	M	B	M
	B	B	B

En la tabla 2.1 se aprecia el resultadoal combinar el impacto y la evaluación del parámetro para determinar el RL, las cuales tienen declaradas las escalas cualitativas como se aprecia en la tabla 2.2.

Tabla 2.2. Escalas cualitativas seleccionadas de las variables lingüísticas.

Impacto (W)	**Evaluación del parámetro (EP)**
A: alto	B: Bien
M: medio	---
B: bajo	M: Mal

La variable impacto está definida según la metodología Magerit 3.0,como el daño sobre el activo derivado de la materialización de la amenaza(M.H.A.P., 2012b).

Laevaluación del parámetroes la valoración de la configuración correcta de la seguridad del parámetro i en el servidor monitoreado.

Cada parámetro i de la lista de chequeo tiene un impacto y una evaluación cualitativa.

La evaluación cualitativa de los parámetros se realiza a través de las formas de evaluación identificadas a continuación.

- **Las formas de evaluación de los parámetros**

Se gestionó el conocimiento tácito y explícito. Como se ha planteado existe mucho conocimiento y experiencias de los especialistas que no se ha documentado de ahí la necesidad imperiosa de gestionar sobre todo el conocimiento tácito. Basado en el resultado de este proceso se establecen las formas de evaluación de los parámetros a partir de los valores obtenidos de las configuraciones de seguridad recopiladas de los SGBD. Así se determinaron tres formas de evaluar los parámetros de las listas de chequeo de seguridad del CIS para los SGBD, las que se reflejan a continuación.

I. Forma de evaluación porlista de términos con comportamiento:

Esta forma de evaluación consta de una lista de términos, que pueden ser nominal, numéricos o una combinación de ambos, según requiera el caso, que constituyen la fuente de datos. Un comportamiento, que define la forma en que se buscará un vínculo entre el conocimiento existente en la base de conocimientoy el resultado recibido del monitoreo proveniente del componente anterior. De esta forma de evaluar, van a existir dos valoraciones, una que se asigna en el caso de que el vínculo resulte positivo y otra para el caso contrario. Esta forma de evaluación está preparada para recibir como resultado un conjunto de datos. Los tipos de comportamientos definidos son:

- **Primera coincidencia:** Se busca la existencia de un mismo valor tanto en el conjunto de datos guardado para un parámetro i, como en la matriz creada para el

diagnóstico. La búsqueda se detiene en cuanto se encuentre al menos un valoren la fuente de datos.

- **Conteo y comparación:** Cuenta la cantidad de valoresprovenientesde la matriz de diagnósticopara un parámetro y realiza una comparación de igualdad con los existentes en elconjunto de datos. Para utilizar esta forma de evaluar, es necesario que los valores sean numéricos.
- **El conjunto de datos es un subconjunto de los valores de la matriz de diagnóstico:**Consiste en verifica si cada uno de los términos pertenecientes al primer conjunto pertenecen también al segundo.
- **El conjunto de valores de la matriz de diagnóstico es un subconjunto del conjunto de datos de la base de datos:**Consiste en verifica si cada uno de los valores de pertenecientes al primer conjunto pertenecen también al segundo.
- **Diferencia de dos a dos:** Este comportamiento verifica que cada término del resultado se encuentre sólo una vez en dicho conjunto. El conjunto de datos para las formas de evaluación con este comportamiento ha de ser vacío, pues no se interactúa con el mismo.

II. Forma de evaluación porvariantes:

Esta forma de evaluación consta de una lista de varianteso de soluciones posibles (no tienen que ser números), cada una de las cuales tiene una evaluación asociada. El resultado se asume como un término único y no como un conjunto, el cual se compara con cada una de las variantes registradas y en caso de encontrar la igualdad, se devuelve la evaluación asociada a dicha variante. Existe una evaluación por defecto, que es opcional, para el caso de que ninguna de las variantes conocidas sea igual que el resultado.

III. Forma de evaluación porintervalos:

Esta forma de evaluación es muy parecida a la anterior, la diferencia radica en que lo que existen son intervalos de números reales y el resultado debe ser un término numérico, del cual se verifica su pertenencia a los diferentes intervalos registrados. Al lograr confirmar la pertenencia se procede a devolver la evaluación correspondiente al intervalo.

Las formas de evaluar no son fijas, en este componente se tiene previsto que se le pueda adicionar o modificar las formas de evaluar según las tres clasificaciones anteriormente vistas. Esta característica va a permitir darle flexibilidad a las formas de evaluar para que puedan adaptarse y evaluar los distintos valores de los parámetros que pueden aparecer en las auditorías.

En la figura 2.5 se muestra la secuencia de pasos de este componente.

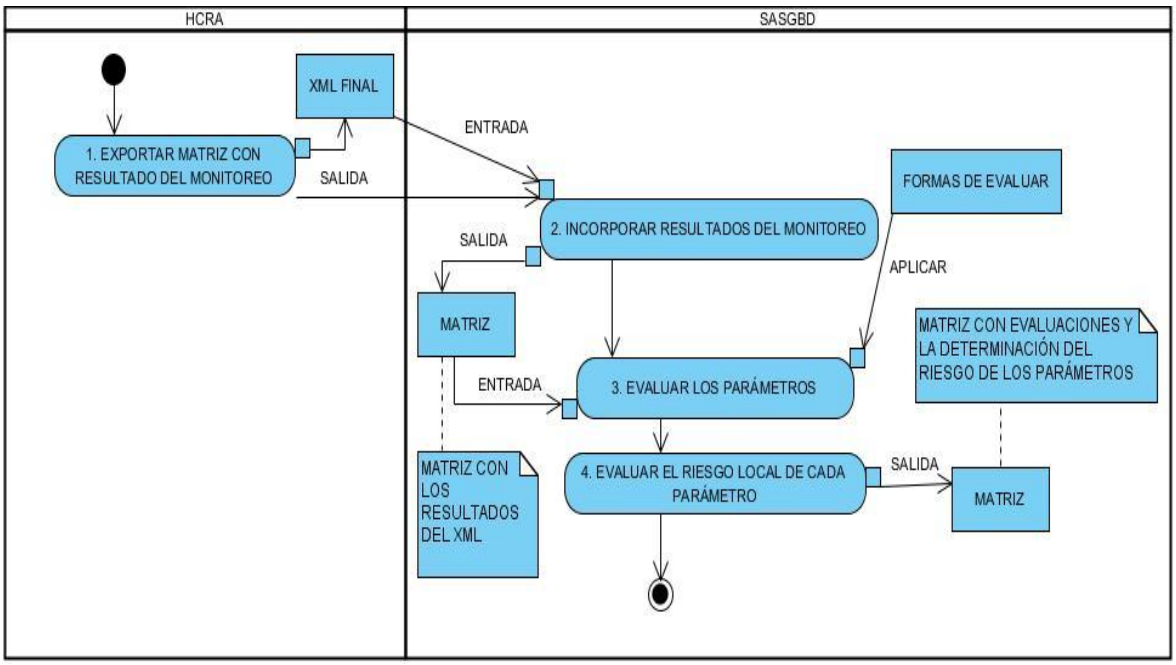


Figura 2.5.Secuencia del componente Estimación del riesgo local.

A pesar de lograr especificar varias formas de evaluar los valores de los parámetros, no todos los parámetros pueden quedar evaluados, por lo que el resultado final de esta variable,quedaen manos del experto, como paso final. En caso de que el auditor deje algún parámetro por evaluar, se toma como valor por defecto del RL el Bajo.

c) Componente: Elección inteligente del RSI

Con este componente se determina el RSI y para la estimación del resultado, este componente se apoya en el uso de las técnicas de la IA. El mismo utiliza como entrada la salida del componente anterior: la evaluación del RLde cada parámetrode la lista de chequeo de seguridad y se convierte en un caso a diagnosticar.

El diseño de los casos es la siguiente: Losrasgos predictores van a estar constituidos por la evaluación del riesgo de los parámetros (RL) de la lista de chequeo de seguridad del CIS para los SGBD (Ver anexo 3).Losrasgos predictoresvaríanpara cada lista de chequeo asociada a unSGBD diferente.

En el anexo 3 se encuentran los parámetros de la lista de chequeo para el servidor Microsoft SQL Server 2000.En el caso del SGBD PostgreSQL, la lista de chequeo fue confeccionada por los propios expertos del Departamento de Seguridad Informática de ETECSA debido a su inexistencia en el sitio oficial del CIS. La elaboración de esta lista es guiada por las mismas pautas que el resto de las listas de chequeo de seguridad publicada por el CIS("C.I.S.", 2013b).

Teniendo en cuenta las diferencias entre las listas de chequeo de seguridad del CIS, se crea la BCteniendo en cuenta una política de particionamiento, como fueron seguidas en los algoritmos WCOID(Smiti, A. y Elouedi, Z., 2010)y COID (Smiti, A. y Elouedi, Z., 2014).

La figura 2.6, ilustra el modo en el cual está estructurada la BC. Como se aprecia, la misma se organizó en forma de árbol, lo cual permitió que los casos que corresponden a un SGBD, no estén relacionados con los de otro SGBD.

A su vez, existen otros niveles o ramas, antes de llegar a los casos como es la versión y la evaluación del rasgo objetivo. La forma de organizar los casos, favorece el acceso y la recuperación de los casos. La variable n en la figura 2.6 se utiliza para transmitir el significado de un número finito de SGBD, de versiones o de casos.

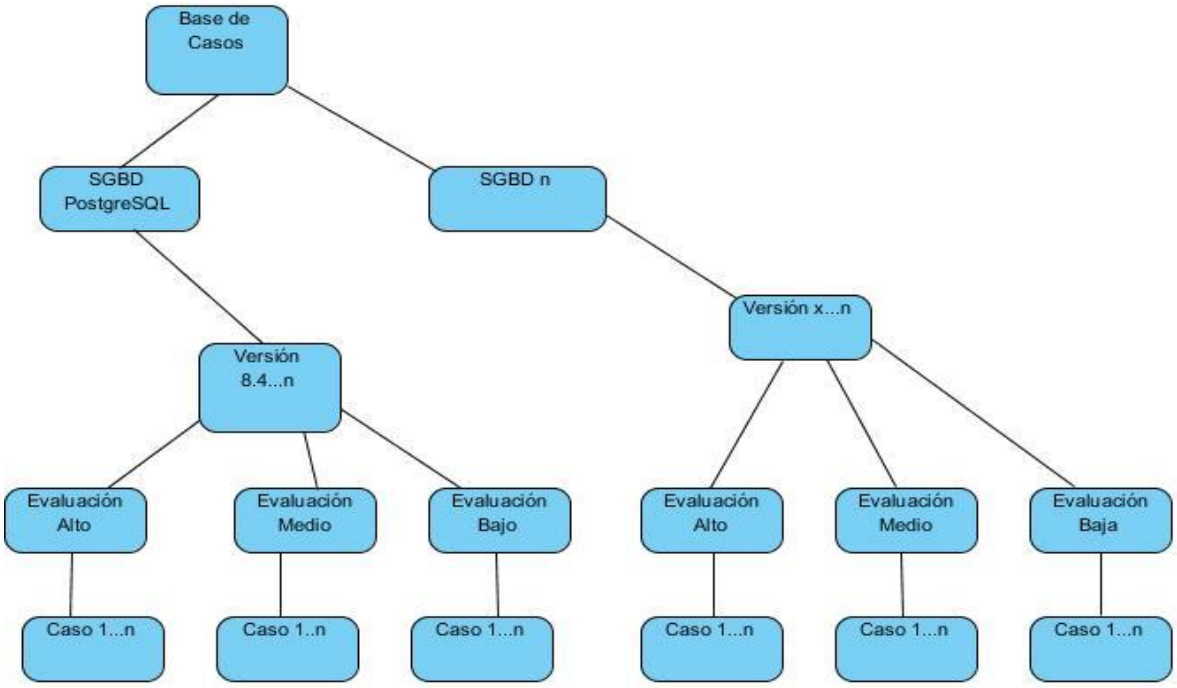


Figura 2.6. Representación de la estructura de la BC.

El rasgo objetivo no es más que la evaluación del riesgo de seguridad de la información (RSI) para un determinado SGBD, el cual tiene en su dominio, los valores Alto, Medio y Bajo.

En este componente, para determinar el número difuso del rasgo objetivo (R), se basa en la propuesta publicada en Patra y Mondal(2015) con la incorporación de algunas modificaciones. Las causas se deben en lo siguiente: la obtención en el componente anterior del RL, por lo que se sustituye en la ecuación por (impacto $\otimes$ PE), para simplificar el proceso de cómputo sin afectar la calidad del resultado. Se incorpora una nueva variable (el peso) para ajustar el valor del RL. Además se realizó una comparación entre la propuesta del autor (ecuación 2.1) y la publicada en Patra y Mondal(2015).En las tablas 2.3, 2.4 se evidencian los resultados obtenidos con la fórmula existente en Patra y Mondal(2015) y con la propuesta del autor, utilizando los casos de estudio definidos en el epígrafe 3.4.1:

- En los casos O_1 yO_2, los resultados son superiores con la propuesta del autor en los valores de semejanza(utilizando la ecuación 2.5), aunque proporcionan la misma evaluación.
- En el caso O_4 los resultados de semejanza no son los esperados pero son parecidos entre los valores de ambas propuestas, pero mejores en la propuesta del autor debido a que en la fórmula propuesta en Patra y Mondal(2015), el segundo valor de semejanza es para el valor lingüístico Bajo, cuando debe ser el Alto como sucede en la propuesta hecha por el autor.
- En los casos del O_5 alO_8 los resultados del diagnóstico es el mismo, pero en la propuesta del autor el valor de semejanza es mayor.
- En los casos O_9 yO_{10} los resultados proporcionan un mismo diagnóstico, pero las semejanzas obtenidas entre el Medio y el Bajo en Patra y Mondal(2015) se acercan más a lo esperado como se especifica en el diseño de la misma.

Se le realizaron pruebas de normalidad a los resultados de semejanza como se aprecia en el anexo 11; las cuales determinaron que para la variable Alto, los datos presentan una distribución normal y no presentan una distribución normal para las variables Medio y Bajo, en ambas propuestas.

En el anexo 12 se encuentra el resultado de la prueba paramétrica T para los datos obtenidos de la variable Alto y para las variables Medio y Bajo, la prueba no paramétrica de los rangos con signo de Wilcoxon. El resultado de aplicar estas pruebas evidenció que:

- Para la variable Alto resultó que la tabla de las muestras relacionadas indica que no hay diferencias estadísticamente significativas entre la propuesta de Patra y Mondal y la del autor.
- Para la variable Medio resultó que existen diferencias significativas con una media mayor para la propuesta de Patra y Mondal.

- Para la variable Bajo resultó que existen diferencias significativas con una media mayor para la propuesta del autor.

La cantidad de casos de estudio con resultados favorables a la propuesta del autorsegún el análisis descrito anteriormente. Además los resultados de las pruebas estadísticas de los valores de semejanza no permiten determinar una propuesta de las analizadas, por lo que se selecciona la fórmula 2.1 para determinar el número difuso del rasgo objetivo (R).

Tabla 2.3. Tabla de resultados de R con diferentes ecuaciones.

Casos de estudio	**R de Patra y Mondal**(2015)	**R de la propuesta del autor**
O_1	(0; 0; 0.3; 0.55; 1)	(0; 0; 0.3; 0.36; 1)
O_2	(0.489; 0.733; 0.986; 0.990; 1)	(0.6; 0.66; 1; 1; 1)
O_3	---	(0.31; 0.37; 0.59; 0.65; 1)
O_4	(0.216; 0.318;0.635; 0.759; 1)	(0.317; 0.349; 0.670; 0.698; 1)
O_5	(0.202; 0.306; 0.590; 0.732; 1)	(0.237; 0.261; 0.576; 0.612; 1)
O_6	(0.166; 0.253; 0.534; 0.697; 1)	(0.189; 0.208; 0.521; 0.562; 1)
O_7	(0.156; 0.237; 0.521; 0.690; 1)	(0.186; 0.205; 0.518; 0.559; 1)
O_8	(0.200;0.301;0.577;0.725; 1)	(0.233;0.256;0.572; 0.608; 1)
O_9	(0.084;0.130;0.407;0.624; 1)	(0.092; 0.101; 0.407; 0.458; 1)
O_{10}	(0.030;0.046;0.350;0.584; 1)	(0.028;0.031;0.330;0.389; 1)

Tabla 2.4. Resultados de semejanza R con diferentes ecuaciones.

Casos de estudio	**Semejanza de R de Patra y Mondal(2015)**	**Semejanza de R de la propuesta del autor**
O_1	Alto: 0.095 Medio: 0.467 Bajo:**0.844**	Alto: 0.028 Medio: 0.339 Bajo:**0.964**
O_2	Alto: **0.854** Medio: 0.392 Bajo: 0.065	Alto: **1.0** Medio: 0.320 Bajo: 0.028
O_3	No es posible evaluar de Medio por esta vía	Alto: 0.320 Medio: **0.957** Bajo: 0.339
O_4	Alto: 0.353 Medio: **0.834** Bajo: 0.360	Alto: 0.363 Medio: **0.883** Bajo: 0.310
O_5	Alto: 0.325 Medio: **0.847** Bajo:0.386	Alto: 0.257 Medio: **0.823** Bajo: 0.428

O_6	Alto: 0.278 Medio: **0.774** Bajo: 0.452	Alto: 0.208 Medio: **0.694** Bajo: 0.517
O_7	Alto: 0.267 Medio: **0.755** Bajo: 0.471	Alto: 0.206 Medio: **0.688** Bajo: 0.521
O_8	Alto: 0.318 Medio: **0.837** Bajo: 0.392	Alto: 0.252 Medio: **0.811** Bajo: 0.436
O_9	Alto: 0.180 Medio: 0.606 Bajo: **0.626**	Alto: 0.115 Medio: 0.481 Bajo: **0.729**
O_{10}	Alto: 0.127 Medio: 0.516 Bajo: **0.751**	Alto: 0.062 Medio: 0.387 Bajo: **0.883**

La ecuación propuesta por el autor es de la siguiente manera:

$$R = (\sum_{i=1}^{n} peso_i \otimes RL_i) ø (\sum_{i=1}^{n} peso_i) (2.1)$$

En la anteriorecuación, lavariableR representa el riesgodel servidor con un número difuso trapezoidal generalizadoR= (t1, t2, t3, t4; w).Son números reales t1, t2, t3, t4 y w tal que: 0 ≤ t1 ≤ t2 ≤ t3 ≤ t4 ≤ 1, 0 ≤ w ≤ 1. La variable w representa la altura del trapecio.

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

1) $\mu_R$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) $\mu_R(x) = 0$ donded $\leq x < \infty$;

Es importante utilizar en esta investigación el peso como medida de distinción entre los parámetros o rasgos que componen un caso, debido a que entre rasgos con igual nivel de

RL, pero con diferencias más determinantes, en unos más que en otros, para la evaluación del riesgo. Por tanto, la utilización del peso contribuye a la diferenciación y en la precisión de la estimación del RSI. En el anexo 3 se encuentran los pesos para los parámetros pertenecientes al SGBD Microsoft SQL Server 2000. Se determinó a través de la escala de Likert, la cual fue aplicada utilizando la encuesta ubicada en el anexo 4.El valor del peso está definido en el intervalo [0,1].

Los valores de la variableRL obtenidosdelos rasgospredictoreso parámetros, son valores cualitativos que se les asociannúmeros difusostrapezoidales según la variable cualitativa, los cuales son creados a partir de los expertos con la aplicación de la encuestadel anexo 4. En la tabla 2.5 se puede apreciar los números difusos a utilizar en la ecuación, según los valores cualitativos de cada parámetro i que posea cada SGBD en una auditoría. En la figura 2.7 se puede observar la representación gráfica de los números difusos trapezoidales generalizados de los valores lingüísticos creada con el MATLAB 7.6.0.

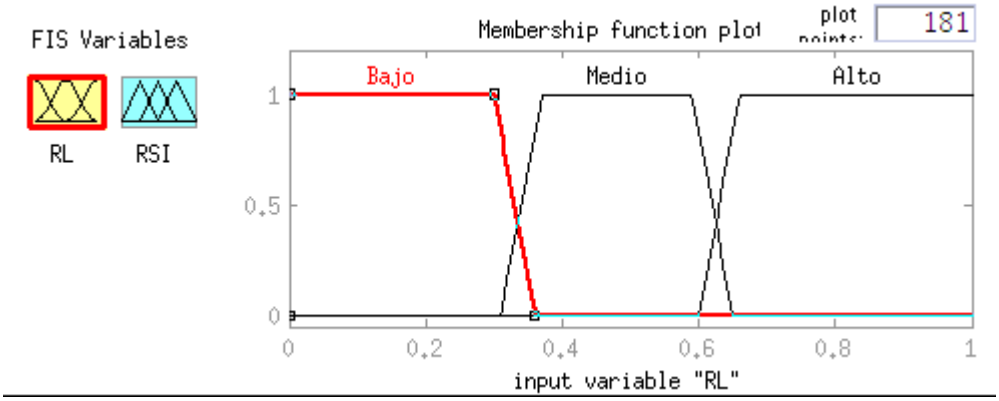


Figura 2.7. Representacióngráfica de los números difusos trapezoidales generalizados de los valores lingüísticos.

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

$$\mu(x) = \begin{cases} 0, \text{ si } x \leq t_1 \\ \frac{(x-t_1)}{t_2-t_1}, \text{ si } t_1 < x \leq t_2 \\ 1, \text{ si } t_2 < x \leq t_3 \\ \frac{(t_4-x)}{t_4-t_3}, \text{si } t_3 < x \leq t_4 \\ 0, \text{ si } x > t_4 \end{cases} \quad (2.2)$$

La función de pertenencia o miembro Gamma es de la siguiente forma:

$$\mu(x) = \begin{cases} 0, \text{ si } x \leq t_1 \\ \frac{(x-t_1)}{t_2-t_1}, \text{ si } t_1 < x \leq t_2 \\ 1, \text{ si } x > t_2 \end{cases} \quad (2.3)$$

La función de pertenencia o miembro L es de la siguiente forma:

$$\mu(x) = \begin{cases} 1, \text{ si } x \leq t_3 \\ \frac{(t_4-x)}{t_4-t_3}, \text{ si } t_3 < x \leq t_4 \\ 0, \text{ si } x > t_4 \end{cases} \quad (2.4)$$

Función de semejanza

La propuesta general de función de semejanza seleccionada es la publicada en Vicente, Mateos y Jiménez (2013):

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

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

En otro caso:

$$S(RN, RO) = 1 - \left(\frac{1-\alpha-\beta}{2} + \alpha\right)\frac{\sum|t_{RNi} - t_{ROi}|}{4} - \left(\frac{1-\alpha-\beta}{2} + \beta\right)|X_{RN} - X_{RO}| \quad (2.5)$$

Donde la variableRN 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 variableRO se corresponde al riesgo expresado en un número difuso trapezoidalde un caso

almacenado en la BC. La función $S(RN, RO)$determina la semejanzaentre los casos. El valor 1 representa la exacta similitud entre los casos, $\alpha + \beta < 1$, $\mu(R)$ es la función miembro del número difusoR.

$$\beta l_{\infty}[(X_{RN}, Y_{RN}), (X_{RO}, Y_{RO})] = \underset{\alpha+\beta<1}{máx}((|X_{RN} - X_{RO}|), (|Y_{RN} - Y_{RO}|)),\ \mu_{RN\cap RO}(x) = \underset{[0\leq x\leq 1]}{mín}(\mu_{RN}(x),\ \mu_{RO}(x)),\ \mu_{RN\cup RO}(x) = \underset{[0\leq x\leq 1]}{máx}(\mu_{RN}(x), \mu_{RO}(x)) \quad (2.6)$$

(X_{RN}, Y_{RN}) y (X_{RO}, Y_{RO}) son los centroides de RN y RO y secalculan de la siguiente manera(Vicente, E., Mateos, A., y Jiménez, A., 2013):

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

Las variablesα y $\beta = {}^{1}/_{3}$ para que se puedan comparar los resultados con el análisis dispuesto enVicente, Mateos y Jiménez (2013).

En caso de no encontrar un caso similar en la base de conocimiento (BC), se tomará como opción la comparación con los números difusos de los valores lingüísticos de la tabla 2.5, en la cual w=1.Estos números difusos son creados en acuerdo con los auditores encuestados y teniendo en cuenta los números utilizados por Vicente, Mateos y Jiménez (2013). Los mismos pueden ser reajustados según las necesidades del auditor que trabaje con el modelo, procurando la flexibilidad del mismo. En el anexo 4 se puede encontrar la encuesta aplicada para determinar los números difusos trapezoidales.

Tabla 2.5. Representación difusa de los valores 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)

Se utiliza el RBCcomo técnica de IA para expresar las auditorías pasadas como conocimiento a utilizar en el diagnóstico de futuras auditorías, necesidad expresada en la encuesta diagnóstico aplicada del anexo 1.

Componente recuperación.**Umbral de semejanza**

El umbral de semejanza es necesario para lograr llegar a determinar el nivel de semejanza entre los casos. Es ampliamente aceptado que la selección de un correcto umbral,conllevaa la extracciónde buenos objetos(Chaira, T. y Ray, A. K., 2004).

Para la definición del umbral de semejanzaexisten varias formulacionespara números difusos (Ruíz , J. S., Alba , E. C., y Lazo , M. C., 1995), (Martínez, N. S., García, M. M. L., y García, Z. Z., 2009)y(Chaira, T. y Ray, A. K., 2004). Se selecciona lasiguiente ecuaciónutilizadapor(Martínez, N. S. y otros, 2009)por su probada utilidad para el RBC y para el cálculo del umbral utilizando objetos en conjuntos difusos(Montellano, J. J. B. y Ruíz, J. S., 1994).

$$\beta_0 = \frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} S(RN, RO) \qquad (2.7)$$

Donde:

β_0: es el valor del umbral de semejanza,$\beta_0 \in L$ y L= [0,1]

m: Número de casos en la BC

i: Valor que recorrerá las filas

j: Valor que recorrerá las columnas

El umbral obtenido,es almacenado en la base de datos de la aplicación SASGBD, para no volver a repetir los cálculos y ahorrar tiempo en el cómputo cada vez que se necesiten este valor. Solo será necesario recalcularlos, en el caso de la entrada de un nuevo caso o por la acción contraria, la eliminación para el mantenimiento de la BC.

Componente recuperación. **Algoritmo de indexación de los casos**

Para la indexación hay algoritmos que se ocupan en parte en la clasificación para una rápida recuperación. El nombre del sistema gestor de la base de datos y la versión se utilizan para aplicar el primer paso en la clasificación de los casos. Los casos almacenados tienen en su rasgo objetivo un valor de tres posibles, el cual permite la clasificación de los casos en grupos más reducidos. Al utilizar estas informaciones, se creó un mecanismo más directo para la indexación de los casos, al tomarla misma intención de la política de particionamiento del algoritmo de indexado de Wong, Mauricio y Papa(2014).

Algoritmo: Indexar_caso(CasoActual, GestorBD, VersionBD) {

1. Identificar la rama del árbol jerárquico de la BC se debe utilizar según el gestor y la versión de la base de datos auditada.

   Si los valores de GestorBDy VersionBDson iguales anombre del gestor de los casos guardadosy de la versión de los casos guardados {
2. Tomar la posición de los casos que pertenecen a la rama de árbol jerárquico en la BC con el mismo gestor y versión.
3. Buscar en la rama identificada, los casos según la evaluación dada para clasificar el objeto CasoActual a indexar

   Si CasoActual.evaluacion igual a “Alto” o igual a “Medio” o igual a “Bajo” {
4. Adicionar el objeto CasoActual de forma jerárquica como una hoja de la subrama Evaluación, según el valor de la evaluación del riesgo del caso nuevo.

}}

5. Si no existe la rama que cumpla con el nombre del gestor o con la versión o la evaluación, se debe crear una nueva rama con el mismo nombre del gestor dado (GestorBD), con una subrama para la versión (VersionBD) y con la subrama con la evaluación del nuevo caso (CasoActual.evaluacion) como se muestra en la figura 2 e insertar como nueva hoja el caso nuevo (CasoActual).

}//Cierre del algoritmo

En la tesis doctoral (Herrera, J. a. Q. y otros, 2013) se utiliza el algoritmo de indexación B-Trees(Bayer, R. y Mccreight, E., 1972), presente en las base de datos relacionales para la recuperación de los casos. Con este algoritmo se solventa en esta investigación la organización de los casos de forma jerárquica en la indexación de los casos. Alejando al algoritmo propuesto, de la organización de los casos.

Componente recuperación.**Algoritmo de acceso y recuperación de los casos**

El algoritmo de acceso y recuperación tiene una secuencia lógica de pasos(Gutiérrez, I. M. y Bello, R. E. P., 2003). Basado en esta misma descripción se propone un método apoyado por el método de indexación descrito anteriormente.

Los casos se almacenan en una base de datos referencial, el cual posibilita la recuperación de los casos a través de consultas SQL. Como variables de entrada, son importantes el nombre (GestorBD) y la versión del sistema gestor de bases de datos (VersionBD), premisas indispensables del algoritmo.Las variablesanteriores son importantes porque la BC está organizadasegún estos datos.

El algoritmo tiene como característica la entrega de una lista de casossemejantes con respecto al caso analizado.

Para agilizar el procesamiento de este algoritmo, se utilizauno de los pasosdelalgoritmo de agrupamiento el K-means,referenciado por Rui y Wunsch(2009), el cual está basado en la función objetivo. Esto se debe a que los casos están agrupados, entre otras características, según el valor de su rasgo objetivo.Elpropósito es determinarel centroide o el casorepresentante de cada grupo de casos bajo el mismo valor del rasgo objetivo. Este paso tomado fue basado en el trabajo deMartínez,Ferreira y García(2008) donde se tiene una base de casos con una estructura jerárquica y se emplea el RBC. Para determinar los

centroides se utiliza la función de semejanzaespecificada por la presencia de números difusos trapezoidales.

$$\text{Centroide} = \sum_{i=1}^{K} \sum_{RO \in C_i} (S(RM_i, RO))^2 \qquad (2.8)$$

Donde:

RM_i: Se corresponde al riesgo expresado en un número difuso trapezoidal de un caso almacenado en la BC y catalogado como el centroide del grupo i.

RO: Se corresponde al riesgo expresado en un número difuso trapezoidalde un caso en un grupo ialmacenado en la BC.

K: Es la cantidad de grupos, K = 1porque la búsqueda del centroide es para un solo grupo conocido.

C_i: Es el conjunto de casos almacenados en un mismo grupo i.

Algoritmo: ObtenerCasosRepresentantes (Lista_base_conocimiento) {

1. Crear lista L2
2. Guardar en lista L1, grupo de casos clasificados con la misma evaluación
3. Buscar dentro de un ciclo, en la variable L1 el centroide con la función de semejanza definida y utilizar la fórmula del umbral para determinar el nivel de semejanza
4. Guardar en lista de casos representantes (L2) el centroide encontrado
5. Volver al paso 2 en caso de encontrarse otro grupo de casos, sino ir al paso 6
6. Devolver L2}

Aparte de buscar los casos representantes de cada grupo por el cual está clasificado los casos según la evaluación, se toman para el análisis de las recomendaciones, los 10 primeros casos más semejantes. Se tiene en cuenta que si no te controla la cantidad de casos semejantes a utilizar en el componente adaptación, esto puede repercutir negativamente en el tiempo de respuesta. Para dar flexibilidad al modelo, se puede variar esta cantidad fijada de número decasos semejantes para analizar.

Algoritmo: AccesoRecuperación(CasoActual, GestorBD, VersionBD,cantCasos=10){

1. Obtener la lista de casos representantes(L1) = ObtenerCasosRepresentantes (Lista_base_conocimiento) Si L1 no está vacía{
    2. Ciclo para buscar el caso más semejante utilizando L1 yel umbral de semejanzaalmacenado o calculado.
    3. Obtener una lista de casos (L2) de tamaño cantCasosque pertenecen al mismo grupo que el caso representante más semejante.
    4. Devolver la lista de casos semejantes.

}}

Componente reutilización. **Algoritmo deadaptación de los casos**

El método de adaptación utiliza el primer caso de la lista de casos semejante que devuelve el algoritmo anterior, para evaluar el SGBD auditado. En esta lista,todos los casos poseen la misma evaluación del RSI.

El método de adaptación utilizado es nulodesde el punto de vista del tipo de transformación al valor del rasgo objetivo, debido a que no hay necesidad de hacerle ningún tipo detransformación.

Se emplea la lista de casos semejantes creada por el método de acceso y recuperación para el diagnóstico de caso y enriquecer el nuevo caso con las recomendaciones.

En caso de no encontrar al menos un caso semejante con la cual proporcionar la evaluación del RSI al SGBD, se evalúa comparando el número difuso del riesgo del servidor por la vía publicada por Patra y Mondal(2015). Solución que no ofrecenAbbasianjahromiy Rajaie(2013), quienes no especificancuál decisión tomar en los casos donde no se encuentre similitud con algún caso. Con esta alternativa no es posible lograr enriquecer los rasgos del caso nuevo con recomendaciones, pero puede ser la entrada a un nuevo caso en la BC. Sugerencia que se le realiza al experto, quien toma la decisión final de su

incorporaciónyde la evaluación final del caso nuevo. La decisión de utilizar como primera opción el RBC en este algoritmo se debe a los resultados que se obtienen en los epígrafes 3.6.1y 3.6.2.

Algoritmo: AdaptaciónCaso(CasoActual,GestorBD, VersionBD**)** {

1. Obtener la lista de casos semejantes(L1)con el algoritmo AccesoRecuperación (CasoActual, GestorBD, VersionBD,10)
2. Si L1 no está vacía{
    3. Evaluar el caso nuevo a diagnosticar utilizando la evaluación del primer caso de la lista.
    4. Ciclo para tomar cada rasgo t del caso CasoActual.
        5. Ciclo anidado para enriquecer con recomendaciones los rasgos del caso CasoActual, utilizando los rasgoscon el mismo valor del riesgo y con la misma evaluación del parámetro de los rasgos existentes de los casos semejantes encontrados en la lista L1.

    }
6. Si L1 está vacía {
    7. Evaluar el caso o el objeto CasoActual, a través de la comparación con los números difusos de los valores lingüísticos utilizados en la evaluación del riesgo.
    8. Recomendar al auditor el nuevo caso evaluado como posible caso a incorporar a la BC.

        }

    }
9. Devolver el caso nuevo evaluado (CasoActual)

}//Cierre del algoritmo

La evaluación final del servidor auditado es responsabilidad del experto, así como de las recomendacionespropuestas por el algoritmo, para cada parámetro o rasgo perteneciente al caso evaluado.

Componente revisión de la evaluación del RSI.

Para validar la evaluación del RSI de un caso realizada por elalgoritmo, intervienen un grupo de expertos que se reúnen para diagnosticar la auditoría que se esté ejecutando, los cuales corroborano no el resultado y finalizanla auditoría al SGBD.

Componente retención. **Algoritmo para el aprendizaje del sistema basado en casos**

El autoaprendizaje se logra con la incorporación de los nuevos problemas que se solucionan a través de la decisión del experto. Para el autoaprendizaje se toma el valor que se obtiene de la fórmula de semejanza. Si la semejanzasobrepasa el umbral, se toman como un caso totalmente nuevo y se introduce en la BC.

Para que el aprendizaje sea efectivo, debe ser de un modo ordenado. Por eso es importante en esta etapa, asegurar una correcta ubicación, la cual define el tiempo de respuesta. Por eso, se utiliza dentro del algoritmo de aprendizaje, el de indexación.

La entrada de un nuevo caso a la BC afecta los valores de las variablespor otros algoritmos que no han tenido en cuenta el nuevo caso incorporado como son: el umbral para la semejanza, el centroide o caso representativo, la utilidady el TM del grupo asociado al mismo caso a incorporar a la BC. Por lo que seprecisa en el método de aprendizaje, reajustar las variables mencionadas y mejorar la respuesta de solución, con cada nuevo conocimiento incorporado.

Algoritmo:Aprendizaje(CasoActual, GestorBD, VersionBD){

1. Indexar_caso(CasoActual, GestorBD, VersionBD)

2. Calcular umbralpara la semejanza, teniendo en cuenta el nuevo caso (CasoActual)
3. CalcularCasoRepresentante con la lista_base_conocimiento y con la misma evaluación.

}// Cierre del algoritmo

La evaluación final de la auditoría

Por último, el componente debe terminar con la evaluación de la auditoría de seguridad informática, donde se debe reflejar el resultado de la evaluación de RSI. Para esto, es necesario tener en cuenta, cuántos servidores hospedan el SGBD que se quiere auditar y efectuar el mismo procedimiento con todos para obtener la evaluación del RSI. Este tipo de situación sucede, cuando la aplicación informática que se audita contiene más de servidor donde tiene hospedado su base de datos.

Una vez obtenida laevaluación delRSI por cada servidor, se tienen en cuenta un conjunto de reglas para determinar la evaluación del RSI de la auditoría de seguridad informática para los SGBD auditados, plasmándose en el informe redactado por el auditor o el equipo al frente de la auditoría.

Las reglas son:

- Si al menos un servidor es evaluado de Alto entonces la evaluación en el informe es Alto.
- Si al menos un servidor es evaluado de Medio y ningún otro evaluado de Alto entonces la evaluación en el informe es Medio.
- Si al menos un servidor es evaluado de Bajo y ningún otro evaluado de Alto y Medio entonces la evaluación en el informe es Bajo.

Las reglas plasmadas son extraídas en consenso con los expertos en auditorías de seguridad informática para los SGBD.Para determinar las reglas, se aplicó la encuesta ubicada en el anexo 5.

2.3 Funcionamiento del modelo propuesto

El modelo para la evaluación del RSIpara los SGBD en la figura 2.1, ofrece un conjunto de componentes que se relacionan entre sí.

Como entrada del modelo se requiere una de las listas de chequeo del CIS orientada a la revisión de la configuración de seguridad informática, con la cual se va a monitorear el SGBD. Para seleccionar la lista de chequeo se debe tener en cuenta el tipo deSGBD y su versión instalada.

Se debe contar con la presencia del administrador del servidor de bases de datos para responder preguntas del auditor con respecto a las políticas de seguridad de la entidad auditada y de la configuración del servidor.

Es necesario que el administrador del SGBD permita el acceso al auditor o al menos que realice por él mismo, las tareas necesarias que debe efectuar el auditor, para el monitoreo y sustracción de los datos de configuración de la seguridad del servidor.

El modelo cuenta con una primera fase: el Monitoreo, donde se ejecuta el componente Obtención de la configuración.Esta fase se realiza en la entidad donde se encuentra hospedado el SGBD. La finalidad es extraer las configuraciones de seguridadactuales en el servidor auditado.

Existen un conjunto de herramientas que automatizan el proceso de monitoreo de la seguridad del SGBD. Se recomienda utilizar la soluciónde sarrollada para este componente, el HCRA. La misma facilita el funcionamiento del siguiente componente del modelo, la cual prepara los datos extraídos del servidor para la entrada a la fase Diagnóstico.

El componente Estimación del RL,tiene lugar en la fase Diagnóstico, en el componente que leprosigue. La misma se efectúa fuera de la entidad donde se realizó el monitoreo del SGBD, preferentementeen el área de trabajo de los auditores. El objetivo de este componente primeramente es la determinación de forma cualitativa de las variables evaluación del parámetro y del riesgo de cada parámetro de la lista de chequeo de seguridad. Para determinar la primera variable se utilizan las formas de evaluar especificadas en este componente. Se utiliza de la metodología Magerit 3.0, una adaptación de la técnica análisis de la evaluación por tablas, cual define como alcanzar el riesgo de cada parámetro.El componente tiene entre otras funciones,incorporar nuevas formas de evaluar, a medida que el auditor pueda identificarlas así como nuevos parámetros que pueden ser incorporados para buscar nuevas vulnerabilidades para un SGBD.

Losvalores cualitativos de la evaluación del RL de cada parámetro, son la entradapara el componente Elección inteligente del RSI, el cuales el encargado finalmente de llegar al resultado final de la evaluación.La respuesta esfundamentada de forma cualitativa en: Alto, Medio o Bajo. Se apoya en la función de similitud incorporada al RBC, la cual permite utilizar números difusos. De este modotomar las experiencias en auditorías anteriores y manejar los términos ambiguos o imprecisos y proporcionar la respuesta cualitativa perteneciente al resultado final del diagnóstico.En caso de no encontrarse un caso similar que pueda determinar la evaluación del nuevo caso, como alternativa, se aplica la función de semejanza entre el número difuso del caso a evaluar y los números difusos asociados para cada variable lingüística de la tabla 2.5 y de esta manera encontrar una evaluación del caso.

2.4 Indicaciones metodológicas para el empleo del modelo propuesto

Para la utilización del modelo propuesto para el análisis del RSI en los SGBD, es necesario tener en cuenta varios aspectos del modelo para su correcto funcionamiento.Dentro de estas se debe tener presente la identificación de la lista de chequeo de seguridad de la CIS ("C.I.S.", 2013b) que corresponda con el SGBD a auditar. Además un personal preparado en las formas de evaluación establecidas en esta investigación y en la técnica de análisis de tabla de la metodología Magerit 3.0 0(M.H.A.P., 2012a)con las modificaciones realizadasen este trabajo.

Inicialización del modelo

Se debe iniciar el modelo por la fase de monitoreo y utilizar la aplicación HCRA, creada para automatizar el proceso llevado a cabo en esta fase y estandarizar los datos de configuración de seguridad que se sustraen del servidor de SGBD que se quiere auditar.Existen varias herramientas que permiten realizar esta tarea como se hace referencia en el capítulo 1, pero la salida de la aplicación mencionada, está preparada para que pueda ser utilizada por el SASGBD y de esta forma permitirle al auditor entrar en la siguiente fase del modelo con los datos preparados para que puedan ser manipulados por el sistema.

En la fase de diagnóstico, el auditor debe tener presente que para el buen funcionamiento del sistema experto SASGBD, es necesario que el mismo termine y corrija las evaluaciones de los parámetros proporcionados por este este sistema, de la matriz cargada con los datos de la configuración obtenida en la fase anterior.

Zona de configuración

Otro aspecto que se debe tener en cuenta por el auditor, es la zona de configuración del sistema experto, donde existen un conjunto de variables necesarias para el sistema. Estas variables repercuten directamente en el resultado final de la evaluación. Los valores de las

variables pueden ser suministradas por el experto, pero el sistema tiene valores predeterminados. Estas variables son: cantidad de casos a contener en la lista de casos a devolver por el algoritmo de recuperación, los valores de los números difusos asociados a las variables lingüísticas y el valor del umbral de semejanza difuso.

El resultado del RSI

El resultado de la evaluación de la matriz de diagnóstico, puede verse manifestada de dos formas, según como se ejecute el componente inteligente del modelo. En caso de que la solución sea dada a través del RBC, puede tener la evaluación, una propuesta de recomendaciones para algunos o todos los parámetros de la matriz de diagnóstico, añadiéndole las posibles recomendaciones especificadas por el auditor o experto.De este modo el auditor puede analizar y fundamentar el resultado final.

Si el resultado es dado a través de la lógica difusa, el enriquecimiento de la respuesta con las recomendaciones sugeridas no es posible, por lo que el experto, debe concluir esta parte en su totalidad. El sistema muestra el grado de semejanza calculado para secundar el resultado arrojado por el sistema y que le permita al auditor tener una fundamentación en la cual apoyarse para su propio diagnóstico.

El modelo tiene como objetivo apoyar al auditor en seguridad informática en la evaluación del riesgo del SGBD y llegar a dar posibles recomendaciones para minimizar los riesgos de cada parámetro de la lista de chequeo de seguridad. Lo que brinda como resultado final es un riesgo potencial; así se le denomina a la medida del daño probable sobre un sistema(Públicas, M. D. H. Y. A., 2012). El riesgo residual que queda dado por un cierto conjunto de salvaguardas desplegadas y una medida de la madurez del proceso de gestión(Públicas, M. D. H. Y. A., 2012), queda fuera del alcance del modelo planteado en esta investigación. Tampoco es objetivo sustituir ninguna de las metodologías de

evaluación del RSI revisadas, sino apoyarse en una de estas, para el funcionamiento del modelo.

2.5 Conclusiones parciales

En el diagnóstico realizado se evidencianlas limitaciones y problemas en los auditores encuestadospara la realización de las auditorías de seguridad informática, en cuanto a la exactitud y el tiempo de respuesta en la evaluación del RSI.

El modelo propuesto contiene los módulos que conforman el proceso de auditoría de seguridad informática y establece un conjunto deprincipios, componentes, estructuración e indicaciones metodológicas para su implementación como una vía de solución a las principales dificultades detectadas en el diagnóstico.

El modelo está basado en el conocimiento y la lógica difusa para la evaluación del RSI a los SGBD, lo cual favorece la reutilización de los resultados de auditorías de seguridad informática anteriores y manejar las ambigüedades de los términos lingüísticos para ajustar las evaluaciones y presentar una mayor consistencia en los resultados. De esta manera contribuir al aumento de la exactitud en la evaluación del RSI en los SGBD y disminuir el tiempo de respuesta.

CAPÍTULO 3.
VALIDACIÓN Y APLICACIÓN DE LOS RESULTADOS

CAPÍTULO3.VALIDACIÓNDEL MODELO PROPUESTO

En el capítulo se exponenlos resultados de la validación del modelo para la evaluación del RSI, a partir del empleo de métodos científicos, que posteriormente son triangulados. Se presentan los resultados de la valoración de los expertos sobre la contribución del modelo en la solución del problema de investigación. Se valora la satisfacción de los usuarios y la aplicabilidad del modelo a través de la aplicación de la técnica de Iadov. Se analizan los resultados del estudio de casos y del experimento donde se aplicó el modelo y los sistemas informáticos desarrollados.

3.1 Proceso de validación de los resultados

Se propone un proceso de validación la cual incluye métodos científicos cuantitativos y cualitativos para comprobar la hipótesis de la investigación. La figura 3.1ilustra el esquema general de este proceso.

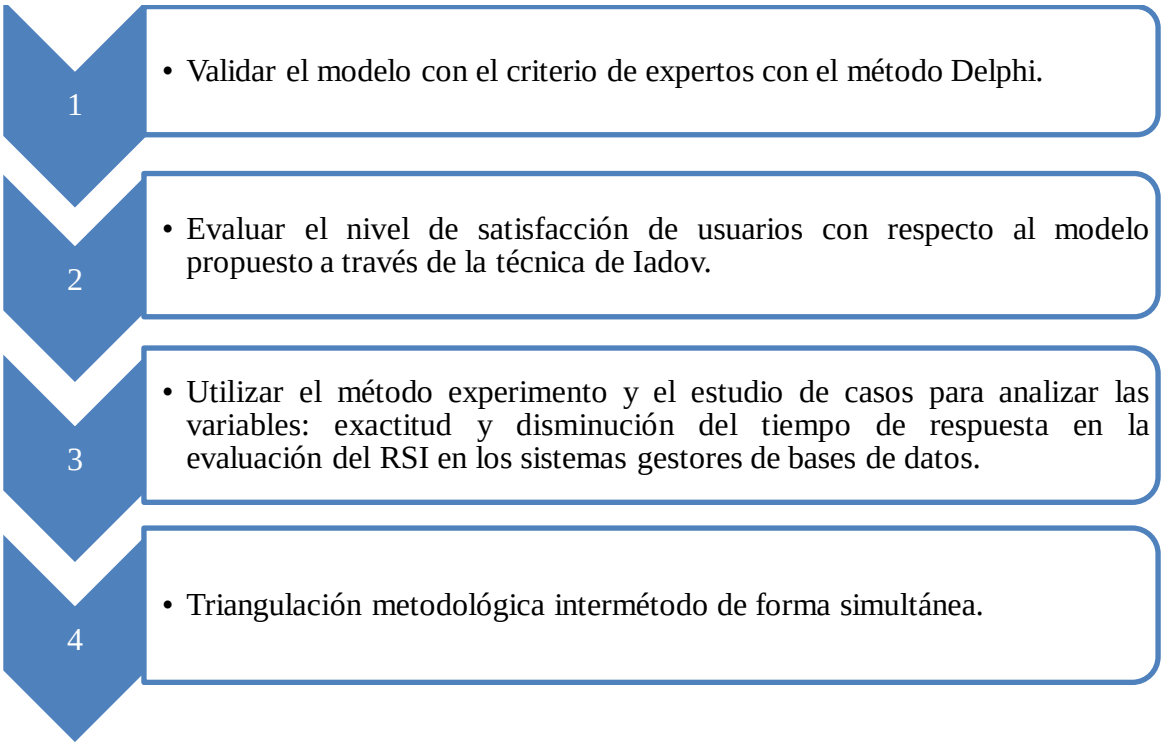


Figura 3.1:Esquema general del proceso de validación.

3.2 Descripción de los resultados de laaplicación del método Delphi

El fundamento Delphi se basa en el análisis de las ideas respecto del futuro de un grupo de expertos en un área del conocimiento, encaminada a la búsqueda de un consenso de opiniones (Rodríguez, J. M. P., Aldana, L. V., y Villalobos, N. H., 2010).

Para la aplicación de esta metodología se apoyó en el cuestionario ubicadoen el anexo 6. El objetivo es conocer el nivel de acuerdo o desacuerdo con las características de los componentes, estructurae integralidad del modelo para la evaluación del RSI en los SGBD.

Para la selección de los expertos se tomacomo base los auditores de seguridad informática de la misma entidad donde se realizó el diagnóstico inicial

La competencia de los expertos en el tema se determina de acuerdo con su opinión sobre el nivel de conocimiento acerca del problema que resuelve y con las fuentes de argumentación que validan sus criterios a través del coeficiente de concordancia de Kendall. Este permite establecer el grado de concordancia de los expertos en relación con todos los aspectos evaluados(González, C. G. y Felpeto, C. G. G. C. G. a. B., 2006). En el anexo 7 se muestra los criterios para seleccionar el experto, así como el resultado de la aplicación del coeficiente de concordancia de Kendall.

Como resultado se seleccionaron expertos de la lista de especialistas presentes en la tabla del anexo 8. Los marcados en negrita, fueron los seleccionados.

De los 33 posibles candidatos, solo 17 respondieron el cuestionario, obteniéndose de estos8con un coeficiente de competencia alto y 8 de medio. La cantidad de expertos necesarios es de 11 como se muestra en la tabla 3.2, por lo que para completar la cantidad se seleccionaron tres expertos con coeficiente de competencia medio, pero próximos al alto, K mayor e iguala 0.75. En la Tabla 3.1 se muestra un resumen de la composición del panel de expertos que participaron en la consulta.

Tabla 3.1. Composición del grupo de expertos

Composición	**Cantidad**
Formación académica universitaria	10
Formación académica técnico medio	1
Más de 20 años de experiencia	1
De 10 a 20 años	3
De 5 a 10 años	3
Menos de 5 años	4

Para determinar el tamaño el número de expertos, se recurrió a la siguiente fórmula publicada en (Legrá, A. a. L., 2014):

$$n = p * q * (\frac{Z_{1-\frac{a}{2}}}{d})^2 \quad (3.1)$$

Donde:

n: es el número de expertos.

d:es el error admisible (cuando d tiende a 0 el número n aumenta).Se expresa como un valor en "tanto por uno" (o sea, porcentual dividido por 100).Y se asume del 50% de p(Legrá, A. a. L., 2014).

p y q:es la proporción o probabilidad de fallo al escoger el experto (su valor está entre 0 y 1); se toma 0.5 cuando se desconoce su valor.

q:=1-p. Cuando p=0,5 se obtiene el mayor valor de n para αy d conocidos.

α: nivel de significación, $\alpha \in [0; 1]$,

El valor de$Z_{1-\frac{a}{2}}$para algunos valores de α son los siguientes:

Tabla 3.2.Valores para calcular la cantidad de expertos (Legrá, A. a. L., 2014).

Nivel de confianza	**α**	**$1-\frac{a}{2}$**	**Z**	**n**

90%	0.10	0.950	1.64	$0.5 * 0{,}5 * \left(\frac{1.64}{0.25}\right)^2 = 10.7 \approx 11$

El resultado del método de consulta de expertoevidenció que tanto los elementos teóricos, las características y las funciones de los componentes del modelo, tienen una alta valoración positiva por parte de los expertos.

La evaluación del modelo como un todo así como de su estructura fue del 91% entre excelente y muy de acuerdo, lo que significa el alto grado de aprobación que tiene el mismo.Se puede apreciar en más detalle los resultados de la aplicación de este método en el anexo 9 y en la figura 3.2.

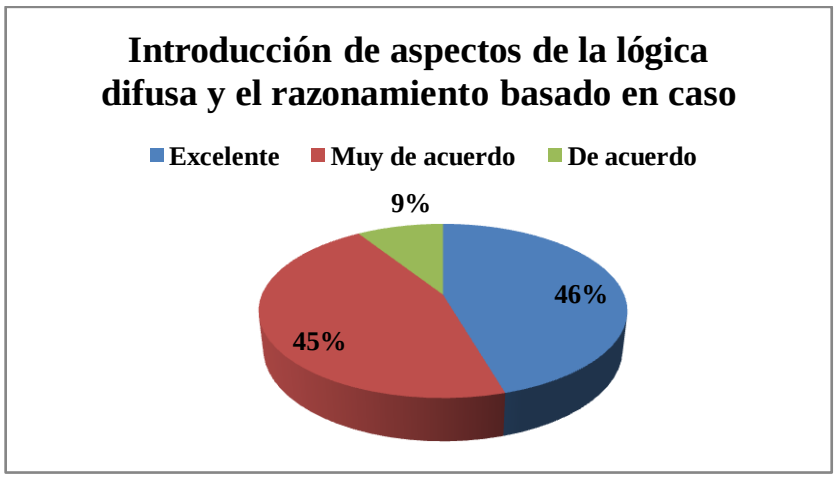


Figura 3.2:Criterios de los expertos sobre el modelo propuesto.

3.3 Evaluar el nivel de satisfacción de usuarios con respecto al modelo

Se midió el nivel de satisfacción del usuario con respecto al modelo implementado por las soluciones informáticas HCRAy SASGBD,para la evaluación del RSI en los SGBD.En aras de cumplir esta meta, se aplicó la técnica de Iadov(Kuzmina, N., 1970)originalmente creada para el estudio de la satisfacción por la profesión en carreras pedagógicas y utilizada para medir la satisfacción del usuario en investigaciones (Cañizares, R. G., Estrada, V. S., y Febles, J. P., 2012; Febles, O. D. y otros, 2012). Se aplicaesta técnica a

través de la encuesta delanexo 6, la cual permite conocer el grado de satisfacción sobre el sistema implementado. La relación entre las preguntas cerradas se establece a través del denominado "Cuadro Lógico de Iadov" como se muestra en la tabla 3.3.

Tabla 3.3: Cuadro lógico de Iadov.

	9. ¿Considera usted que se deba evaluar el riesgo de seguridad de la información en los sistemas gestores de bases de datos sin un modelo coherente?								
	No			No sé			Sí		
11. ¿Le satisface este modelo para la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos?	**10. ¿Si usted necesitara realizar la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos usaría este modelo?**								
	Sí	No sé	No	Sí	No sé	No	Sí	No sé	No
Me gusta mucho	1	2	6	2	2	6	6	6	6
No me gusta tanto	2	2	3	2	3	3	6	3	6
Me da lo mismo	3	3	3	3	3	3	3	3	3
Me disgusta más de lo que me gusta	6	3	6	3	4	4	3	4	4
No me gusta nada	6	6	6	6	4	4	6	4	5
No sé qué decir									

El número resultante de la interrelación de las tres preguntas nos indica la posición de cada sujeto en la escala de satisfacción.

Para medir el grado de satisfacción de los usuarios respecto al modelo para la evaluación del RSI en los SGBD, se toma una muestra de 15 usuarios de 33 posibles, del Departamento de Seguridad Informáticade ETECSA quienes completaron estas preguntas del cuestionario. Entidad conla cual se estableció el compromiso de entrega de la solución informática.

El resultado de la satisfacción individual fue el siguiente:

Tabla 3.4. Resultado individual de la técnica de Iadov.

Escala	**Resultado**	**Cantidad**	**%**
1	Máximo de satisfacción	10	67
2	Más satisfecho que insatisfecho	3	20

3	No definida	2	13
4	Más insatisfecho que satisfecho	0	0
5	Clara insatisfacción	0	0
6	Contradictoria	0	0
Total		15	100

Para determinar el índice de satisfacción grupal (ISG) se trabaja con los diferentes niveles de satisfacción que se expresan en la escala numérica que oscila entre +1 y -1 de la siguiente forma:

Tabla 3.5. Escala de calificación del nivel de satisfacción

+1	Máximo de satisfacción
0,5	Más satisfecho que insatisfecho
0	No definido y contradictorio
- 0,5	Más insatisfecho que satisfecho
-1	Máxima insatisfacción

El índice de satisfacción grupal (ISG) se calcula por la siguiente fórmula:

$$\text{ISG} = \frac{10\ (+1) + 3\ (+0{,}5) + 2\ (0) + 0\ (-0{,}5) + 0\ (-1)}{17} = \frac{11{,}5}{15} = 0{,}76666 \approx 0{,}77 \qquad (3.2)$$

Las variables: A, B, C, D, E, representan el número de sujetos con índice individual 1; 2; 3 ó 6; 4; 5 y N representa el número total de sujetos del grupo.

Los valores de ISG que se encuentran comprendidos entre -1 y -0.5 indican insatisfacción; los comprendidos entre -0.49 y $+0.49$ evidencian contradicción y los que caen entre 0.5 y 1 indican que existe satisfacción.

El índice de satisfacción grupal calculado es0,77 aproximadamente lo que significa una clara satisfacción con la propuesta y reconocimiento de su utilidad en una mayor exactitud y disminución del tiempo de respuesta en la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos.

3.4 Estudio de casos

El método de estudio de caso es una herramienta valiosa de investigación(Martínez Carazo, P. C., 2011). Es seleccionada esta metodología por el incremento en la actualidad para las pruebas de hipótesis, comola meta mayor de los investigadores de los sistemas de información de hoy (Dubé, L. y Paré, G., 2003).Debido a la carencia o la inexistencia de bases de datos internacionales con las cuales se pueda contrastar los resultados y el nivel secreto de los datos que se le confieren a las auditorías de seguridad informáticas realizadas a las entidades.

En el estudio de caso no se selecciona una muestra representativa de una población sino una muestra teórica(Martínez Carazo, P. C., 2011)y adicionar casos hasta la saturación de la teoría. Se argumenta que el número de casos apropiado depende del conocimiento existente, del tema y de la información que se pueda obtener a través de la incorporación de estudios de casos adicionales y los casos elegidos debe satisfacer los criterios de selección establecidos por el investigador en forma previa. No existe un rango o número determinado de la muestra, sin embargo algunos sugieren entre cuatro y diez casos si no hay un número ideal (Schank, R., 1982).

Se considera utilizar los casos múltiples como una herramienta poderosa para crear teoría porque permiten la replicación y la extensión entrecasos individuales(Martínez Carazo, P. C., 2011).

Los criterios de selección de los casos de estudio especificados por el autor son los siguientes:

- Casos que describan escenarios de auditorías de seguridad de informáticas a SGBD en situaciones extremas que pudieran existir (O_1, O_2 y O_3).
- Un caso que describan un escenario de auditorías de seguridad de informáticas a SGBD con todos los parámetros con riesgo local Alto y los pesos con valor 1 (O_4).

- Casos que describan escenarios de auditorías de seguridad de informáticas a SGBD que alternen parámetros con impacto Alto, donde existan pesos con valor 1 y evaluados con riesgoAlto (O_5,O_6,O_7,O_8,O_{11},O_{12},O_{13},O_{14},O_{15}).
- Casos que describan escenarios de auditorías de seguridad de informáticas a SGBD que alternen parámetros evaluados con riesgo local enAlto, Medioy Bajo y los parámetros no tengan peso igual a uno (O_9,O_{10}).

3.4.1 Diseño de casos de estudio

En este apartado se diseñó unestudio de casos múltiples para comprobar la exactitud de la evaluación del modelo utilizando como lista de chequeo, la ubicada en el anexo 3.Los casosvan a estar identificado para el SGBD Microsoft SQL Server 2000. Los casos de estudios fueron proporcionados por los especialistas encuestados, tomados de escenarios reales para este SGBD.

Los casos de estudio son los siguientes:

- El caso O_1consiste en una matriz a la cual se le asignaa todos los valores de la variable RL de los parámetros de la lista de chequeo de Bajo. El resultado final del riesgo debe ser Bajo.

- El caso O_2 consiste en una matriz a la cual se le asigna a todos los valores de la variable RL de los parámetros de la lista de chequeo de Alto. El resultado final del riesgo debe ser Alto.

- El caso O_3 consiste en una matriz a la cual se le asigna a todos los valores de la variable RL de los parámetros de la lista de chequeo de Medio. El resultado final del riesgo debe ser Medio.

- El casoO_4consiste en una matriz a la cual se le asigna a todos los valores de la variable RLde bajo a todos los parámetros, exceptuando aquellos cuyos pesos fuesen igual a uno. A estos parámetros se les evaluó de Alto. Los parámetros son: 1.4, 1.8, 1.9, 2.2, 2.11, 2.12, 2,13, 2.15, 3.1, 3.2, 3.4, y 3.5.Estos parámetros son muy críticos y pueden determinar en su conjunto el resultado de una auditoría. El resultado final del riesgo debe ser Alto.

- El diseño de loscasos desO_5hasta el O_{15} se puede encontrar en el anexo 10. Los mismos fueron facilitados por especialistas y tomados de auditorías reales. El resultado final del riesgo de los casos del O_5 hasta el O_8debe ser Alto, pero también son aceptables resultados de medio.El resultado final del riesgo de los casos del O_9, el O_{10}y O_{11} deben ser Bajo y con proximidad al Medio, es decir no tan bajo, el caso O_{10}debe ser más bajo que el anterior. El resto de los valores de los casos del 11 al 15 se encuentra en el anexo 10.

3.5 Losexperimentos

Se definió un conjunto de experimentos para probar la contribución del modelo propuesto a través de las aplicaciones informáticas SASGBD y HCRAen la investigacióna partir del estudio de casos.Las variables que intervienen en el experimento son:

- Un modelo basado en el conocimiento y la lógica difusa (Variable independiente VI1). La variable tiene dos dimensiones, una dimensión cuando está auxiliadade una BC. Otra dimensión cuando no está presente una BC que la respalde.
- La exactitud en la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos (Variable dependiente VD1). El Vocabulario Internacional de Metrología (VIM) ("J.C.G.M.", 2012) define la exactitud de medida como la proximidad existente entre un valor medido y un valor verdadero

de un mensurando (Armenteros, A. M. R., Balboa, J. L. G., y Mingorance, J. L. M., 2010).

- El tiempo de respuesta para arribar a la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos (Variable dependiente VD2).

El instrumento de mediciónutilizado para la medición en la pre y post-prueba fue el SPSS v13.0.

Se diseñaron tres pre-experimentos del tipo con pre-prueba y post-prueba(Hernández, R. S., Fernández, C. C., y Baptista, P. L., 2006):

1. $\boldsymbol{G\ O_1 x O_2}$

Simbología del diseño experimental:

- G: Grupo dematrices de auditorías de seguridad informática a SGBD.
- X: condición experimental(variable independiente de la hipótesis).
- O:medición de las variables dependientes de la hipótesis (O_1, preprueba y O_2, postprueba).

La aplicación del primerpre-experimento tiene como objetivo comparar los resultados de la evaluacióndel RSI obtenidos por la instancia del modelo a través de la lógica difusa con los proporcionaron los especialistas de los casos diseñados anteriormente. En este experimento se realiza con la variable independiente en la dimensión establecida anteriormente como de ausencia de información en la BC.Fueron utilizados los casos del O_1 al O_{15} para llevar a cabo este experimento.

La aplicación de un segundo experimento tiene la misma finalidad de comparar los resultados de la variable independiente en la dimensión establecida como: presencia de una BC, utilizando el RBC como técnica de IA para el diagnóstico. La BC estuvo conformada

por los casos del O_1 al O_{15}. Los casos de entrada para ser diagnosticados del O_1 al O_{15} de tal forma que si se va a evaluar un caso O_x,la misma se excluye de la BC y después es incorporada para excluir la siguiente, donde $0<x<16$.

La aplicación de un tercer experimento tiene como objetivo comparar el tiempo de respuesta para realizar el diagnóstico antes y después del modelo. Fueron utilizados los experimentosanteriores para la medición del tiempo de respuesta de la variable independiente. Se le añade el tiempo que requiere el primer componente del modelo para obtener la configuración de seguridad del servidor monitoreado y del segundo componente al cargar el resultado del primero. A los casos del O_1 al O_{10}se les medirá el tiempo de respuestasin la existencia de una BC y delO_{11} al O_{15} teniendo como base de caso del O_1 al O_{10}.

3.6 Análisis de los resultados

De acuerdo a los experimentos realizados se alcanzaron los siguientes resultados:

3.6.1 Resultados del experimento uno

Para dejar evidenciado los resultados de primer experimento,la medición de la semejanza con los casos de estudioa través del SASGBD,los resultadosfueron reflejados en la tabla 3.6. La columna R es de la palabra resultado, la variable C de correcto, M de Mal yA de aceptable.

Tabla 3.6. Resultado del valor de semejanza para el experimento uno.

Caso	Semejanza Alto	Semejanza Medio	Semejanza Bajo	Valor observado	Valor esperado	R
O_1	0.028	0.339	**0.964**	Bajo	Bajo	C
O_2	**1.0**	0.320	0.028	Alto	Alto	C
O_3	0.320	**0.957**	0.339	Medio	Medio	C
O_4	0.363	**0.883**	0.310	Medio	Alto	M
O_5	0.257	**0.823**	0.428	Medio	Alto	A
O_6	0.208	**0.694**	0.517	Medio	Alto	A
O_7	0.206	**0.688**	0.521	Medio	Alto	A

O_8	0.252	**0.811**	0.436	Medio	Alto	A
O_9	0.115	0.481	**0.729**	Bajo	Bajo	C
O_{10}	0.062	0.387	**0.883**	Bajo	Bajo	C
O_{11}	0.066	0.388	**0.835**	Bajo	Bajo	C
O_{12}	0.239	**0.781**	0.430	Medio	Alto	M
O_{13}	0.368	**0.878**	0.285	Medio	Alto	M
O_{14}	0.173	**0.606**	0.563	Medio	Alto	M
O_{15}	0.300	**0.900**	0.346	Medio	Alto	M

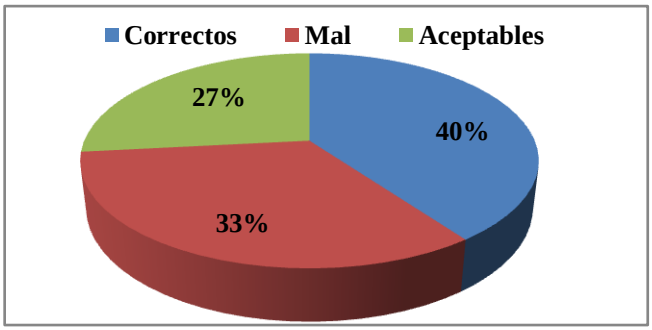


Figura 3.3:Resultados de la medición de la exactitud del primer experimento.

Los resultados de la medición de la exactitud del primer experimento mostrado en la figura 3.3 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.

3.6.2 Resultados del experimento dos

Los resultados del segundo experimento fueron reflejados en la tabla 3.7.La variable US es el umbral de semejanza, CR son los casos representantes determinados para cada caso evaluado para el experimentoy SE de sin evaluar.

Tabla 3.7. Resultado del valor de semejanza para el experimento 2.

Caso	**Semejanza Alto**	**Semejanza Medio**	**Semejanza Bajo**	**Valor observado**	**Valor esperado**	**US**	**CR**	**R**
O_1	0.620	0.500	**0.830**	Bajo	Bajo	0.61	$O_{3,5,11}$	C
O_2	No es capaz de determinar un caso semejante por RBC				Alto	0.68	$O_{3,5,11}$	SE
O_3	**0.833**	---	0.400	Alto	Medio	0.62	$O_{5,11}$	M

O_4	0.786	**0.900**	0.383	Medio	Alto	0.62	$O_{3,5,11}$	M
O_5	0.830	**0.833**	0.510	Medio	Alto	0.61	$O_{3,6,11}$	A
O_6	**0.830**	0.687	0.619	Alto	Alto	0.60	$O_{3,5,11}$	C
O_7	**0.800**	0.660	0.643	Alto	Alto	0.60	$O_{3,5,11}$	C
O_8	0.721	**0.782**	0.543	Medio	Alto	0.61	$O_{3,11,13}$	A
O_9	0.398	0.429	**0.940**	Bajo	Bajo	0.62	$O_{3,11,13}$	C
O_{10}	0.334	0.367	**0.863**	Bajo	Bajo	0.63	$O_{3,9,13}$	C
O_{11}	0.370	0.400	**0.940**	Bajo	Bajo	0.62	$O_{3,9,13}$	C
O_{12}	0.740	**0.810**	0.560	Medio	Alto	0.61	$O_{3,9,13}$	M
O_{13}	**0.935**	0.849	0.377	Alto	Alto	0.61	$O_{4,9,15}$	C
O_{14}	0.557	0.659	**0.716**	Bajo	Alto	0.61	$O_{9,13,15}$	M
O_{15}	**0.849**	0.837	0.462	Alto	Alto	0.61	$O_{9,12,13}$	C

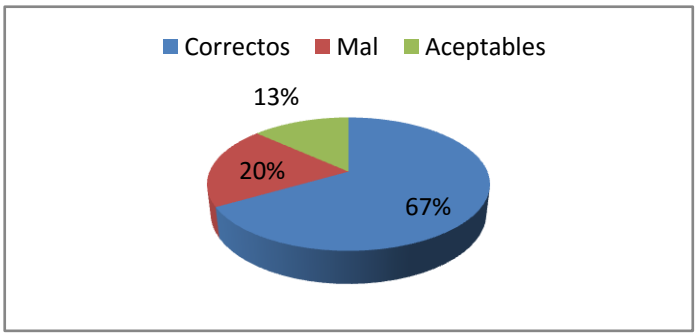


Figura 3.4:Resultados de la medición de la exactitud del segundo experimento.

Los resultados de la medición de la exactitud del segundo experimento,mostrados en la figura 3.4 son que: de los 15 casos, 8 fueron evaluados correctamente, 1 sin evaluar, 4 evaluados incorrectamente y 2 demodo aceptable.

En el resultado para los casosO_{4y} O_{12} se puede apreciar que aunque errónea, se acercan mejor al resultado real porque la semejanza con valor lingüístico Alto es bueno a diferencia del resultado en el primer experimento.

Al comparar los resultados de este experimento con el primero se puedenextraer variasinferencias:

- La técnica RBC tiene mejor resultado en la exactitud en la evaluación del RSI, por eso en el algoritmo de adaptación es la primera opción a utilizar.

- La segunda es que incorporando al algoritmo de adaptación propuesto, la evaluación con la lógica difusa como segunda opción, le da al modelo un mayor

porciento de exactitud. Esta afirmación es posible porque en el experimento dos no se pudo evaluar por el RBC el caso O_2 y no completamente el caso O_3,pero el resultado mostrado en el experimento unopara estos casos fue acertada y se asegura a todos los casos una evaluación, lo cual indica lautilidad de incorporar la lógica difusa en el algoritmo de adaptación.La figura 3.5 muestra el resultado de la medición de la exactitud incluyendo los casos O_2 y O_3 evaluados de forma correcta. Obteniendo 10 casos correctos, ninguno sin evaluar, 2 de manera aceptable y 3 de mal.

- La utilización de la técnica RBC en conjunto con el algoritmo para la obtención de casos representantes, se desempeñan como función de ajuste de los números difusos asociados a los valores lingüísticos de la tabla 2.5. En la tabla 3.7 se puede observar como los casos representativos se van adecuando a medida que la BC se modifica con cada iteración del experimento, por lo cual, indica un reajuste de los casos representativos y por ende de los números difusos con los cuales se realiza la comparación de la semejanza.

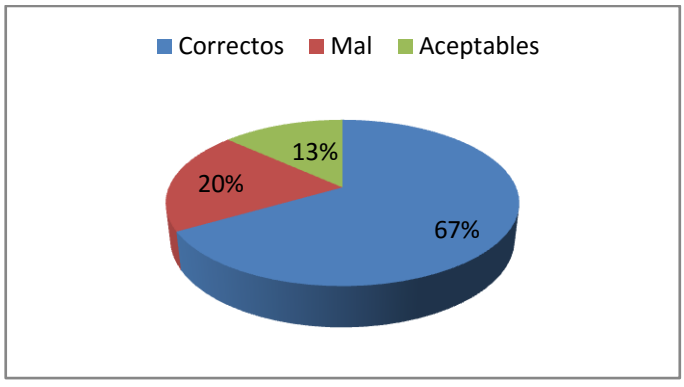


Figura 3.5:Resultados de la medición de la exactitud del modelo.

3.6.3 Resultados del experimento tres

Con la aplicación de este experimento, se evidencia la mejoría en el tiempo de respuesta con la utilización del modelo para contribuir el proceso de auditoría de seguridad

informática a los SGBD. Las mediciones se encuentran en la tabla 3.8. La medición se realizó con un hardware con prestación de RAM de 6 Gb, un procesador Intel Core i3 a 2.53 GHz.

Tabla 3.8. Resultados de la medición para el experimento 3.

Caso	**Tiempo de respuesta con presencia del modelo (TM)**	**Tiempo de respuesta con ausenciadel modelo**	**Veces que mejora (Suponiendo 8 horas.; 8*60/TM)**
O_1	3 minutos	Se demoran horas e incluso días para resolver una sola auditoría (O_x) según respuestas tomadas de la encuesta diagnóstico aplicada del anexo 1.	480/3=160
O_2	3 minutos		160
O_3	3 minutos		160
O_4	3 minutos		160
O_5	4 minutos		120
O_6	4 minutos		120
O_7	4 minutos		120
O_8	3 minutos		160
O_9	3 minutos		160
O_{10}	3 minutos		160
O_{11}	3 minutos		160
O_{12}	3 minutos		160
O_{13}	3 minutos		160
O_{14}	3 minutos		160
O_{15}	3 minutos		160

Con los resultados del tercer pre-experimento no es necesario realizar pruebas no paramétricas ante la evidencia de los resultados positivos del modelo en la contribución de la reducción del tiempo de respuesta y observar la diferencia significativa en los resultados de la medición del tiempo entre la presencia y la ausencia del modelo.

3.7 Triangulación metodológica

A partir de la aplicación de los métodos Delphi, la técnica de Iadov, la aplicación de estudio de casos ylos tres pre-experimentos se realizó la triangulación metodológica de los resultados. El método permite contrastar los resultados obtenidos para determinar las coincidencias y divergencias. Como se afirma en la tesis doctoral(Cañizares, R. G. y otros, 2012), es un procedimiento de control implementado para garantizar la confiabilidad en los

resultados de cualquier investigación. De ese modo se disminuye el sesgo que se produce al comparar los resultados obtenidos en la cuantificación de variables mediante un método cuantitativo y las tendencias y dimensiones que surgen de la aplicación de métodos cualitativos.

En la siguiente tabla se muestran los objetivos y los métodos aplicados.

Tabla 3.9. Objetivos a evaluar con la triangulación metodológica.

Objetivos a evaluar	**Métodos cuantitativos**	**Método cualitativos**
Objetivo 1. La exactitud en la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos. Objetivo 2. El tiempo de respuestaen la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos.	Iadov, El escalamiento de Likert, Encuesta Experimento	Estudio de casos

Luego de aplicada la triangulación metodológica, se obtuvieron los siguientes resultados:

- La valoración satisfactoria por expertos (mediante Delphi)de los componentes, la estructura y de la integralidad del modelo y aceptación favorable de los usuarios (aplicando Iadov y encuestas) del modelo permite concluir que el mismoaumentala exactitud en la evaluación del RSI y la disminución en el tiempo de respuesta del mismo.

- La coincidencia alcanzada entre los resultados de los estudios de casosy pre-experimentos con respecto al efecto de la implementación del modelo, se comprueba la contribución a la solución del problema de investigación.

A partir del análisis de los resultados obtenidos de la aplicación de los métodos tanto cuantitativos como cualitativos, se puede concluir que la hipótesis planteada en la investigación fue confirmada cumpliéndose los objetivos trazados.

3.8 Conclusiones parciales

A partir de la validación realizada:

- Se comprueba la importancia de los componentes que conforman el modelo propuesto, así como sus relaciones y funcionalidades.
- Se demostró la clara satisfacción de los usuarios del modelo con la propuesta desarrollada con la técnica de Iadov.
- Los resultados de lospre-experimentos con casos de estudio evidenciaron la viabilidad del modelo para solventar las variables dependientes planteadas en la investigación.
- La aplicación de la triangulación metodológica inter-métodoevidencia que existe una correspondencia positiva en los resultados que se obtienen.
- El modelo aumenta la exactitud y disminuye el tiempo de respuesta en la evaluación del RSI en los SGBD.

CONCLUSIONES GENERALES

La investigación realizada permite llegar a las siguientes conclusiones:

- A partir de la revisión de los principales referentes teóricos que sustentan la presente investigación, se confirma que los modelos, las listas de chequeo y las metodologías existentes en la literatura, presentan limitaciones para la evaluación del riesgo de seguridad de la información en los SGBD.

- Se fundamentó un modelo basado en el conocimiento y la lógica difusa que contribuye a mejorar la exactitud y el tiempo de respuesta de la evaluación del RSI en los SGBD.

- Con la validación del modelo se constató que el mismo contribuye en mayor medida a la exactitud y la disminución del tiempo de respuesta en la evaluación del riesgo de seguridad de la información.

RECOMENDACIONES

En vista del alcance y limitaciones de la presente investigación, se recomienda:

- Continuar las investigaciones que permitan extender el modelo que incluya la evaluación del riesgo de seguridad de la información teniendo en cuenta las vulnerabilidades del sistema operativo donde se encuentre hospedado el SGBD.

- Valorar la incorporación del SCAP a la herramienta desarrollada como instancia del modelo,por la tendencia actual de crear las listas de chequeo compatibles a este protocolo.

REFERENCIAS BIBLIOGRÁFICAS

1. "A.E.C.". (2013). Extraído desde 04/05/2014, de http://www.aec.es/web/guest/centro-conocimiento/seguridad-de-la-informacion.

2. "C.I.S.". (2013a). Extraído desde 14/11/2013, de http://benchmarks.cisecurity.org/downloads/audit-tools/.

3. "C.I.S.". (2013b). Extraído desde 12/11/2013, de http://benchmarks.cisecurity.org/downloads/multiform/.

4. "C.I.S.". (2014a). Extraído desde 18/02/2014, de http://benchmarks.cisecurity.org/membership/certified/.

5. "C.I.S.". (2014b). Extraído desde 18/02/2014, de http://benchmarks.cisecurity.org/membership/.

6. "E.N.I.S.A.". (2013). Extraído desde 15/12/2013, de http://rm-inv.enisa.europa.eu/comparison.html?menu1=&menu2=&Button=+Go+.

7. "E.N.I.S.A.". (2014). Extraído desde 28/03/2014, de http://rm-inv.enisa.europa.eu/tools/t_octave.html.

8. "E.Y.". (2017). Extraído desde 18/01/2017, de http://www.ey.com/ar/es/services/advisory/cybersecurity.

9. "Forbes". (2017). Extraído desde 18/01/2017, de http://www.forbes.com/companies/ernst-young/.

10. "J.C.G.M.". (2012). International vocabulary of metrology – Basic and general concepts and associated terms (VIM) (3ra ed., pp. 108). (Francia): Oficina Internacional de Pesas y Medidas (BIPM).

11. "Microsoft". (2011). Extraído desde 18/01/2014, de http://www.microsoft.com/business/es-es/Content/Paginas/article.aspx?cbcid=225.

12. "Microsoft". (2014). Extraído desde 18/01/2014, de http://technet.microsoft.com/es-es/library/cc185712.aspx.

13. "N.I.S.T.". (2011). Managing Information Security Risk *Organization, Mission, and Information System View* (pp. 88). Gaithersburg (EUA): National Institute of Standards and Technology NIST.

14. "N.I.S.T.". (2013). Extraído desde 24/3/2014, de http://web.nvd.nist.gov/view/vuln/statistics-results?cves=on&query=&cwe_id=CWE-89&pub_date_start_month=-1&pub_date_start_year=2007&pub_date_end_month=-1&pub_date_end_year=-1&mod_date_start_month=-1&mod_date_start_year=-1&mod_date_end_month=-1&mod_date_end_year=-1&cvss_sev_base=&cvss_av=&cvss_ac=&cvss_au=&cvss_c=&cvss_i=&cvss_a=.

15. "N.I.S.T.". (2014). Extraído desde 19/02/2014, de https://web.nvd.nist.gov/view/ncp/information.

16. "N.I.S.T.". (2017a). Extraído desde 1/1/2017, de http://web.nvd.nist.gov/view/vuln/statistics-results?cves=on&query=&cwe_id=CWE-89&pub_date_start_month=-1&pub_date_start_year=2007&pub_date_end_month=-1&pub_date_end_year=-1&mod_date_start_month=-1&mod_date_start_year=-1&mod_date_end_month=-1&mod_date_end_year=-1&cvss_sev_base=&cvss_av=&cvss_ac=&cvss_au=&cvss_c=&cvss_i=&cvss_a=.

17. "N.I.S.T.". (2017b). Extraído desde 1/1/2017, de http://web.nvd.nist.gov/view/vuln/statistics-results?cves=on&query=&cwe_id=CWE-264&pub_date_start_month=-1&pub_date_start_year=2007&pub_date_end_month=-1&pub_date_end_year=-1&mod_date_start_month=-1&mod_date_start_year=-1&mod_date_end_month=-1&mod_date_end_year=-1&cvss_sev_base=&cvss_av=&cvss_ac=&cvss_au=&cvss_c=&cvss_i=&cvss_a=.

18. "N.I.S.T.". (2104). Extraído desde 19/02/2014, de http://web.nvd.nist.gov/view/ncp/repository?tier=&product=&category=Server&authority=&keyword=.

19. "S.A.N.S.". (2014a). Extraído desde 19/02/2014, de http://www.sans.org/about#__utma=267470038.1420689340.1392837884.1392837884.1392837884.1&__utmb=267470038.17.9.1392839808182&__utmc=267470038&__utmx=-&__utmz=267470038.1392837884.1.1.utmcsr=%28direct%29|utmccn=%28direct%29|utmcmd=%28none%29&__utmv=-&__utmk=242046337.

20. "S.A.N.S.". (2014b). Extraído desde 19/02/2014, de http://it-audit.sans.org/community/whitepapers.

21. "S.I.G.E.A.". (2013a). Extraído desde 18/02/2014, de http://www.gxsgsi.es/caracteristicas/.

22. "S.I.G.E.A.". (2013b). Extraído desde 18/02/2014, de http://www.gxsgsi.es/que-es-gxsgsi/.

23. "S.O.M.A.P.". (2007). Extraído desde 14/12/2013, de http://downloads.sourceforge.net/project/somap/Infosec%20Risk%20Assessment%20Guide/Version%201.0/somap_guide_v1.0.0.pdf?r=http%3A%2F%2Fsourceforge.net%2Fprojects%2Fsomap%2Ffiles%2FInfosec%2520Risk%2520Assessment%2520Guide%2FVersion%25201.0%2F&ts=1387036307&use_mirror=garr.

24. "Softonic". (2014). Extraído desde 18/01/2014, de http://microsoft-baseline-security-analyzer.softonic.com/.

25. "Technet", M. (2014). Extraído desde 18/01/2014, de http://technet.microsoft.com/es-es/security/cc184924.aspx.

26. 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.

27. Areitio, J. B. (2008). Seguridad de la información: redes, informática y sistemas de información Extraído desde 19/10/2013, de http://www.google.com.cu/books?hl=es&lr=&id=_z2GcBD3deYC&oi=fnd&pg=PP1&dq=gesti%C3%B3n+de+riesgo+de+seguridad+inform%C3%A1tica+ISO/IEC&ots=wqpnvGI0Ph&sig=Vyoy_6v-T-IReKYYOAHmjwi2N58&redir_esc=y#v=onepage&q&f=false.

28. Armenteros, A. M. R., Balboa, J. L. G. y Mingorance, J. L. M. (2010). *Error, incertidumbre, precisión y exactitud, términos asociados a la calidad espacial del dato geográfico.* Paper presented at the Catastro: formación, investigación y empresa: Selección de ponencias del I Congreso Internacional sobre catastro unificado y multipropósito, Extraído desde 2/3/2015, de http://coello.ujaen.es/congresos/cicum/ponencias/Cicum2010.2.02_Ruiz_y_otros_Error_incertidumbre_precision.pdf.

29. Bayer, R. y Mccreight, E. (1972). of Large Ordered Indexes Acta Informatica, Vol. 1, Fase. 3, 1972. *Acta Informatica, 173*, 1S9. Extraído desde 24/03/2015, de http://www.informatik.uni-jena.de/dbis/lehre/ws2005/dbs1/Bayer_hist.pdf.

30. Bello, R. (2002). Aplicaciones de la Inteligencia Artificial. *Ediciones de la Noche, Guadalajara, Jalisco, México. ISBN: 970-27-0177-5.*

31. Bello, R. y Morell, C. (2000). Modelos para el razonamiento basado en casos en presencia de rasgos borrosos *Reporte de investigación terminada. Código 001.535 Bell M, CDICT UCLV.*

32. Bello, R. P. (1998). *Solución de problemas bajo incertidumbre*. Universidad Central de Las Villas. Santa Clara.70

33. Bonzano, A. (1998). ISAC: a Case-Based Reasoning System for Aircraft Conflict Resolution. *Trinity College*.

34. 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.

35. Bytes, S. (2014a). Extraído desde 22/01/2014, de http://www.secure-bytes.com/soa.php.

36. Bytes, S. (2014b). Extraído desde 18/01/2014, de http://www.secure-bytes.com/sqa.php.

37. Calder, A. y Watkins, S. G. (2010). Information Security Risk Management for ISO27001/ISO27002(pp. 187). Extraído desde 21/01/2014, de http://books.google.com.cu/books?id=8Ffa1dOFgO4C&printsec=frontcover&dq=iso+27001&hl=es&sa=X&ei=rPTdUvS-AarXyAHz4YDoBQ&ved=0CDkQ6AEwAg#v=onepage&q=iso%2027001&f=false.

38. Cañizares, R. G., Estrada, V. S. y Febles, J. P. (2012). *Repositorio de recursos educativos para las instituciones de educación superior.* Tesis doctoral, Universidad de las Ciencias Informáticas. Extraído desde 08/06/2014.

39. Carvalho, F. D. y Silva, E. M. D. (2006). Cyberwar-Netwar: Security in the Information Age(pp. 159). Extraído desde 21/01/2014, de http://books.google.com.cu/books?id=_kK-__r1EhgC&pg=PA64&dq=Tuar++risk+analysis&hl=es&sa=X&ei=Z6bqUvObFuGNygHU1YDACQ&ved=0CC4Q6AEwAA#v=onepage&q=Tuar&f=false.

40. Citicus. (2014). Extraído desde 22/01/2014, de http://www.citicus.com/oursoftware_softwarecapabilities.asp.

41. Cole, E. (2011). Network Security Bible(pp. 936). Extraído desde 04/12/2013, de http://books.google.com.cu/books?id=Iq8lPbhGRuYC&pg=PT689&dq=:+Operationally+Critical+Threat,+Asset,+and+Vulnerability+Evaluation&hl=es&sa=X&ei=Zo2qUrOCKMv5kQeJrYGgCQ&ved=0CHAQ6AEwCA#v=onepage&q=%3A%20Operationally%20Critical%20Threat%2C%20Asset%2C%20and%20Vulnerability%20Evaluation&f=false.

42. Coles, R. S. y Moulton, R. (2003). Operationalizing IT risk management. *Computers & Security*, *22*, 487-493. Extraído desde 01/05/2014, de http://www.sciencedirect.com/science?_ob=MiamiImageURL&_cid=271887&_user=12364430&_pii=S0167404803006060&_check=y&_origin=article&_zone=toolbar&_coverDate=30-Sep-2003&view=c&originContentFamily=serial&wchp=dGLbVlB-zSkWb&md5=92d215ca6311c8b57ac9e737e01c15fb&pid=1-s2.0-S0167404803006060-main.pdf.

43. Cuadrado, S. R., González, E. F. R., Curbelo, H. H., Luna, Y. C., Casas, G. C. y Gutiérrez, I. M. (2011). Sistema experto basado en casos para el diagnóstico de la hipertensión arterial. *Revista Facultad de Ingeniería Universidad de Antioquia*, 202-213. Extraído desde 10/06/2014, de http://www.scielo.org.co/pdf/rfiua/n60/n60a20.pdf.

44. Chaira, T. y Ray, A. K. (2004). Threshold selection using fuzzy set theory. *Pattern Recognition Letters, 25*, 865–874. Extraído desde 1/12/2016, de.

45. 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.

46. Chou, T.-S. (2011). Information Assurance and Security Technologies for Risk Assessment and Threat Management: Advances Extraído desde 21/01/2014, de http://books.google.com.cu/books?id=2zlES3JYNpoC&pg=PA243&dq=cramm+risk+a

ssessment+tool&hl=es&sa=X&ei=S3_oUrCuI8abrAGe9IG4Cw&ved=0CGwQ6AEwCA#v=onepage&q=cramm%20risk%20assessment%20tool&f=false.

47. Del Valle, A. L. (2007). *Metamodelos de la investigación pedagógica*. 5. La Habana (Cuba).163

48. Dimitrakos, T., Raptis, D., Ritchie, B. y Stølen, K. (2002). Model-based security risk analysis for Web applications: The CORAS approach. *Proceedings of the EuroWeb*. Extraído desde 10/06/2014, de http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.195.6095&rep=rep1&type=pdf.

49. Dubé, L. y Paré, G. (2003). Rigor in information systems positivist case research: current practices, trends, and recommendations. *MIS quarterly*, 597-636. Extraído desde 30/11/2016, de https://pdfs.semanticscholar.org/c9d1/bcdb95aa748940b85508fd7277622f74c0a4.pdf.

50. Dubois, D., Prade, H. y Klement, E. P. (1999). Fuzzy Sets, Logics and Reasoning about Knowledge Extraído desde 13/9/2016, de http://www.google.com.cu/books?hl=es&lr=&id=Fl2nwu7t8msC&oi=fnd&pg=PR9&dq=Klawonn+1999%2BThe+role+of+Similarity++in+Fuzzy+Reasoning.+Fuzzy+Sets,+Logics+and++Reasoning&ots=Tunmb5EpO_&sig=VB-30llOMAKaT2fiW_qV_nouEtk&redir_esc=y#v=onepage&q=grade%20of%20ownership&f=false.

51. Dutta, S. y Bonissone, P. (1991). Integrating Case-Based and Rule-Based Reasoning: the Possibilistic Connection. *In Uncertainty in Artificial Intelligence, P.P. Bonissone, M. Henrion, L.N. Kanal, J.F. Lemmer (editores), North-Holland*, *6*, 281-298.

52. Eloff, J. H., Labuschagne, L. y Badenhorst, K. P. (1993). A comparative framework for risk analysis methods. *Computers & Security*, *12*, 597-603. Extraído desde 31/01/2014, de http://www.sciencedirect.com/science/article/pii/016740489390056B.

53. Febles, O. D., Estrada, V. S. y Febles, J. P. R. (2012). *MIDAC: modelo para el desarrollo de aplicaciones compuestas basadas en arquitecturas orientadas a servicios.* Tesis doctoral, Universidad de las Ciencias Informáticas. Extraído desde 07/06/2013.

54. García, F. G., González, P. a. C. y Sanchez, A. A. (2011). *Introspección en sistemas de planificación basada en casos aplicados a juegos de estrategia.* Tesis de maestría, Universidad Complutense de Madrid. Extraído desde 12/4/2015.

55. García, J. M. (2008). Análisis y control de riesgos de seguridad informática: control adaptativo. Extraído desde 9/01/2014, de http://www.acis.org.co/fileadmin/Revista_105/JMGarcia.pdf.

56. González, C. G. y Felpeto, C. G. G. C. G. a. B. (2006). Tratamiento de datos Extraído desde 13/12/2013, de http://books.google.com.cu/books?id=AhNx24025ZoC.

57. Guida, G. y Tasso, C. (1994). Design and Development of Knowledge- Based Systems. From Life Cycle to Methodology. *John Wiley and Sons Ltd., Ba_ns Lane, Chichester, England.*

58. Gutiérrez, I. M. y Bello, R. E. P. (2003). *Un modelo para la toma de decisiones usando Razonamiento Basado en Casos en condiciones de incertidumbre.* Tesis doctoral, Universidad Central Marta Abreu. Extraído desde 13/10/2014.

59. Gutiérrez, I. M., Bello, R. E. P. y Tellería , A. R. (2002). Un sistema basado en casos para la toma de decisiones en condiciones de intertidumbre. *Revista Investigación Operacional, 23*. Extraído desde 13/10/2014, de http://rev-inv-ope.univ-paris1.fr/files/23202/IO-23202-1.pdf.

60. Hand, D. J. (1997). Construction and Assessment of Classification Rules. . *John Wiley & Sons, Chichester, UK.*

61. Haouchine, M.-K., Chebel-Morello, B. y Zerhouni, N. (2008). *Competence-Preserving Case-Deletion Strategy for Case-Base Maintenance*. Paper presented at the Similarity and Knowledge Discovery in Case-Based Reasoning Workshop. 9th European Conference on Case-Based Reasoning.

62. Harmon, M. (1996). Reinforcement learning: A tutorial.

63. Hejazi, S. R., Doostparast, A. y Hosseini, S. M. (2011). An improved fuzzy risk analysis based on new similarity measures of generalized fuzzy numbers. *Expert Systems with Applications, 38*, 9179-9185. Extraído desde 10/06/2016, de http://www.sciencedirect.com/science/article/pii/S0957417411001217.

64. Hernández, N. D., Yelandy, M. L. y Cuza, B. G. (2013). Modelos causales para la Gestión de Riesgos. *Revista Cubana de Ciencias Informáticas, 7*, 58-74. Extraído desde 09/02/2014, de http://rcci.uci.cu/index.php/rcci/article/view/475/246.

65. Hernández, R. S., Fernández, C. C. y Baptista, P. L. (2006). *Metodología de la investigación* (Cuarta (ed)). McGraw-Hill Companies. Mexico D.F.882

66. Herrera, J. a. Q., Eduardo, J. O., Alfaro, L. C. y Tupac, Y. V. (2013). *Modelo Estocástico a partir de Razonamiento Basado en Casos para la Generación de Series Temporales.* Tesis doctoral, Universidad Nacional de San Agustín. Extraído desde 07/12/2014, de http://dspace.concytec.gob.pe/bitstream/concytec/76/1/herrera_qj.pdf.

67. Hisham, M. H. y Brunil, D. R. (2009). Asset Identification for Security Risk Assessment in Web Applications. *International Journal of Software Engineering, 2*, 53-74. Extraído desde 09/12/2013, de http://www.ijse.org.eg/Content/Vol2/No3/Vol2_No3_4.pdf.

68. Hsieh, C. H. y Chen, S. H. (1999). Similarity of generalized fuzzy number with graded mean integration representation. *Proceedings of the 1999 eighth international fuzzy systems association world congress, 2*, 551-555. Extraído desde 23/10/2016, de http://dev02.dbpia.co.kr/1/04/38/1043852.pdf?article=877863.

69. International, R. (2014a). Extraído desde 22/01/2014, de http://riskwatch.com/secure-watch/.

70. International, R. (2014b). Extraído desde 22/01/2014, de http://riskwatch.com/riskwatch-360/.

71. Jiankang, L. y Bing, J. (2012). The building of decision support system using case-based reasoning techniques and its application in rural surplus labor problem. *International Journal of Digital Content Technology and its Applications, 6*. Extraído desde 13/04/2015, de http://www.aicit.org/JDCTA/ppl/JDCTA1529PPL.pdf.

72. Juárez, J. y J., P. (2005). Inteligencia Artificial:Razonamiento Basado en Casos. *74*. Extraído desde, de.

73. Karabacak, B. y Sogukpinar, I. (2005). ISRAM: information security risk analysis method. *Computers & Security, 24*, 147-159.

74. Kelly, A. E., Lesh, R. A. y Baek, J. Y. (2014). *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching*. Routledge.

75. Kolodner, J. L. (1992). An Introduction to Case-Based Reasoning. *Artificial Intelligence, 6*, 3-34.

76. Kolodner, J. L. (1993). Case-Based Reasoning,. *Morgan Kaufmann Publishers, Inc., San Mateo, CA, 1993*.

77. Kouns, J. y Minoli, D. (2011). Information Technology Risk Management in Enterprise Environments: A Review of Industry Practices and a Practical Guide to Risk Management Teams Extraído desde 12/12/2013, de http://books.google.com.cu/books?id=0D2eM4GQCqgC&pg=PT104&lpg=PT104&dq=CCTA+%28Central+Communication+and+Telecommunication+Agency%29+Risk+Analysis+and+Management+Method&source=bl&ots=C9GyQb7wmU&sig=mQUnevf8X_Cs-NiNbFtwOoODtW8&hl=es&sa=X&ei=QcqoUsa4CIGmyQHU9YGQBQ&ved=0CD4Q6AEwAg#v=onepage&q=CCTA%20%28Central%20Communication%20and%20Telecommunication%20Agency%29%20Risk%20Analysis%20and%20Management%20Method&f=false.

78. Kuzmina, N. (1970). Metódicas investigativas de la actividad pedagógica. *Moscú, Rusia: Editorial Leningrado*.

79. Lawanna, A. y Daengdej, J. (2010). Hybrid Technique and Competence-Preserving Case Deletion Methods for Case Maintenance in Case-Based Reasoning. *International Journal of Engineering Science*. Extraído desde 1/11/2014, de http://www.ijest.info/docs/IJEST10-02-04-21.pdf.

80. Legrá, A. a. L. (2014). Elementos teóricos y prácticos de la investigación científico-tecnológica Edición 0 (Ed.)(pp. 568). Extraído desde 12/12/2016, de https://www.google.com.cu/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja

&uact=8&ved=0ahUKEwjP0Prhp_HQAhXkslQKHQReDUEQFggaMAA&url=http%3A%2F%2Fwww.ismm.edu.cu%2Fwp-content%2Fuploads%2Fdocumentos%2FETPICT-A2L2-1-993.pdf&usg=AFQjCNETU1D1QLI8c9_iP1F69K7U_F0iOQ.

81. Lerner, U. (2002). Hybrid Bayesian Networks for Reasoning about Complex Systems. *CITESEER*. Extraído desde 14/10/2015, de citeseer.nj.nec.com/lerner02hybrid.html.

82. Auth Unpublished Workor. (1998). *Sistemas Basados en el Conocimiento*. Departamento de Ciencia de la Computación, Facultad de Matemática, Física y Computación, Universidad Central "Martha Abreu" de Las Villas. Santa Clara, Cuba.

83. López , R. D. M. (2005). Retrieval, reuse, revision, and retention in case based reasoning. *The Knowledge Engineering Review, Cambridge University Press DOI: 10.1017/S000000000000000 Impreso en el Reino Unido, 00:0*, 1–2.

84. Lozano, L. y Fernández, J. (2006). Razonamiento Basado en Casos: Una Visión General.

85. M.H.A.P. (2012a). MAGERIT – versión 3.0 Metodología de Análisis y Gestión de Riesgos de los Sistemas de Información Vol. 3. PÚBLICAS, M. D. H. Y. A. (Ed.) *Libro III - Guía de Técnicas* (pp. 42). Extraído desde 08/02/2016, de http://administracionelectronica.gob.es/.

86. M.H.A.P. (2012b). MAGERIT – versión 3.0 Metodología de Análisis y Gestión de Riesgos de los Sistemas de Información Vol. 1. PÚBLICAS, M. D. H. Y. A. (Ed.) *Libro I - Método* (pp. 42). Extraído desde 14/12/2015, de http://administracionelectronica.gob.es/.

87. Main, J., Dillon, T. S. y Shiu, S. C. (2001). A tutorial on case based reasoning *Soft computing in case based reasoning* (pp. 1-28): Springer.

88. Martell, V. F. y Zulueta, Y. V. (2014). Modelo para el análisis de riesgos en Líneas de Productos de Software. *Revista Cubana de Ciencias Informáticas, 8*, 82-98. Extraído desde 20/02/2014, de http://rcci.uci.cu/index.php/rcci/article/view/516/252.

89. Martínez Carazo, P. C. (2011). El método de estudio de caso Estrategia metodológica de la investigación científica. *Revista científica Pensamiento y Gestión*. Extraído desde 14/11/2016, de http://rcientificas.uninorte.edu.co/index.php/pensamiento/article/view/3576/2301.

90. Martínez, N. S., Ferreira, G. L., García, M. M. L. y Valdivia, Z. G. (2008). El Razonamiento Basado en Casos en el ámbito de la Enseñanza/Aprendizaje. *Revista de Informática Educativa y Medios Audiovisuales, 5*, 25-32. Extraído desde 11/04/2014, de http://laboratorios.fi.uba.ar/lie/Revista/Articulos/050510/A4mar2008.pdf.

91. Martínez, N. S., García, M. M. L. y García, Z. Z. (2009). Modelo para diseñar sistemas de enseñanza-aprendizaje inteligentes utilizando el razonamiento basado en casos. *Revista Avances en Sistemas e Informática, 6*, 67-78.

92. Mendenhall, W., Beaver, R. J. y Beaver, B. M. (2010). Introducción a la probabilidad y estadística Editores, C. L. (Ed.)Extraído desde 25/04/2014, de http://latinoamerica.cengage.com.

93. Minka, T. P. (2001). Expectation propagation for approximate Bayesian inference. *In the 17th Annual Conference on Uncertainty in AI (UAI)*, 362-369.

94. Mitra, R. y Basak, J. (2005). Methods of Case Adaptation. *International Journal of Intelligent Systems, 20*, 627-645.

95. Montellano, J. J. B. y Ruíz, J. S. (1994). *Agrupaciones en gráficas difusas.* Maestría en Ciencias Matemáticas, Universidad Nacional Autónoma de México. Extraído desde 30/03/2016.

96. Morell, C. a. P., Bello, R. P. y Grau, R. Á. (2005). *Extensiones al razonamiento basado en casos para su aplicación en la planificación de procesos.* Tesis doctoral, Universidad Central "Marta Abreu" de las Villas. Extraído desde 19/10/2014.

97. Ms-Isac, M.-S. I. S. a. a. C. (2010). Extraído desde 14/12/2013, de http://msisac.cisecurity.org/resources/guides/documents/Risk-Management-Guide.pdf#page=3&zoom=auto,179,0.

98. Muñoz, C. R. (2002). Auditoría en sistemas computacionales(pp. 796). Extraído desde 20/05/2013, de http://www.google.com.cu/books?hl=en&lr=&id=3hVDQuxTvxwC&oi=fnd&pg=PR11&dq=%22Auditor%C3%ADa+en+sistemas+computacionales%22&ots=3dQknhDWfd&sig=HswuIz4NiFMo7qCS5cYHj67TQjo&redir_esc=y#v=onepage&q=%22Auditor%C3%ADa%20en%20sistemas%20computacionales%22&f=true.

99. Navarro, M. I. (2011). *Una nueva perspectiva para recuperación en razonamiento basado en casos: mejora de la adecuación del ca so recuperado usando funciones de riesgo.* Tesis doctoral, Universidad de Granada. Extraído desde 20-10-2014.

100. Nayagam, V. L. G. y Sivaraman, G. (2012). A novel similarity measure between generalized fuzzy numbers. *International Journal of Computer Theory and Engineering, 4*, 448-450. Extraído desde 12/06/2016, de http://search.proquest.com/openview/eec49a1256b540bb2851180ce545830e/1?pq-origsite=gscholar.

101. Nikolić, B. y Ružić-Dimitrijević, L. (2009). Risk assessment of information technology systems. *Issues in Informing Science and Information Technology, 6*, 595-615. Extraído desde 08/01/2014, de http://www.masters.donntu.edu.ua/2013/fknt/godla/library/article6.pdf.

102. Ohno-Machado, L. (1996). Medical Applications of Neural Networks: Connectionist Models of Survival. . *PhD dissertation, Stanford University, Department of Computer Science. Report STAN-CS-TR-96-1564.*

103. Palisade. (2014). Extraído desde 26/01/2014, de http://www.palisade-lta.com/risk/.

104. Pandey, S. K. (2012). A Comparative Study of Risk Assessment Methodologies for Information Systems. *Bulletin of Electrical Engineering and Informatics, 1*, 111-122. Extraído desde 12/12/2013, de http://journal.uad.ac.id/index.php/EEI/article/view/131/pdf.

105. 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 https://www.researchgate.net/publication/270293719_Fuzzy_risk_analysis_using_area_and_height_based_similarity_measure_on_generalized_trapezoidal_fuzzy_numbers_and_its_application?ev=srch_pub&_sg=QCmFQLVPV4JBRTdvJpLjX0HBl9f7uScevCBesAbbQpa_9sjjQ0AhD4Jp3DUq1r5h.X21z_vA8tX86dn_YUq5ZAXMapyCk4mB_nDi3qYmTSyy8veR0vMvCm0Kgen36ecxF.4aWc32RyWiO7Css4x4UAWAw28Jonu7MUZpKtcDEhXeE2aD-T8RXjJBpUmrIARBt7.

106. Peltier, T. R. (2001). Information Security Risk Analysis(pp. 296). Extraído desde 20/05/2013, de http://books.google.com.cu/books?id=O0_fO2Xvp98C&pg=PA231&dq=risk+assessment+security+%2B+expert+system&hl=es&sa=X&ei=Tr_qUq3tF6H7yAHc_ICQAg&ved=0CEQQ6AEwAg#v=onepage&q=risk%20assessment%20security%20%2B%20expert%20system&f=false.

107. Piattini, M. G. V. y De Peso, E. N. (2001). Auditoría Informática. Un enfoque práctico Vol. 1. EDITOR, A.-O. G. (Ed.) Extraído desde 20/05/2013, de http://books.google.com.cu/books?id=0agPPQAACAAJ.

108. Piñero, P. Y. P. y García, M. M. L. (2005). *Un modelo para el aprendizaje y la clasificación automática basado en técnicas de softcomputing*. Tesis doctoral, Universidad Central "Marta Abreu" de las Villas. Extraído desde 30/03/2015.

109. Públicas, M. D. H. Y. A. (2012). MAGERIT – versión 3.0 Metodología de Análisis y Gestión de Riesgos de los Sistemas de Información Vol. 1.(pp. 127). Extraído desde, de http://administracionelectronica.gob.es/.

110. Quigley, M. (2008). Encyclopedia of Information Ethics and Security Extraído desde 25/11/2013, de http://books.google.com.cu/books?id=H2VuBddvMLAC&pg=PT432&dq=importance+of+the+audit+of+security+of+systems&hl=en&sa=X&ei=Z5NEUdzANeaV7AaXmoDgCA&ved=0CEcQ6AEwBQ.

111. Quisbert, A. E. V. (2014). Modelo de Sistemas Multi-Agentes para Percibir, Evaluar y Alertar Ex-Antes los Accesos no Autorizados a Repositorios de Base de Datos. *Revista del Postgrado en Informática*, 128. Extraído desde 5/1/2017, de http://www.revistasbolivianas.org.bo/pdf/rpgi/n1/n1_a24.pdf.

112. Ramakanth, D. y Vinod, K. (2011). SQL Injection - Database Attack Revolution And Prevention. *Journal of International Commercial Law and Technology, 6*, 224-231. Extraído desde 23/09/2013, de http://www.jiclt.com/index.php/jiclt/article/view/141/139.

113. Rodríguez, J. M. P., Aldana, L. V. y Villalobos, N. H. (2010). Método Delphi para la identificación de prioridades de ciencia e innovación tecnológica. *Revista Cubana de Medicina Militar, 39*, 214-226. Extraído desde 3/8/2016, de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0138-65572010000300006&nrm=iso.

114. Rodríguez, Y. S., García, M. M. L. y De Baets, B. (2008). *Generalización de la métrica basada en la diferencia de valores (VDM) para variables lingüísticas y su aplicación en sistemas basados en el conocimiento.* Tesis doctoral, Universidad Central "Marta Abreu" de las Villas. Extraído desde 12/01/2014.

115. Rui, X. y Wunsch, D. C. (2009). *Clustering* (ilustrada (ed), Vol. Volumen 10 de IEEE Press Series on Computational Intelligence). IEEE Press. New Jersey.400

116. Ruíz , J. S., Alba , E. C. y Lazo , M. C. (1995). Introducción al reconocimiento de patrones Vol. 51. *ENFOQUE LOGICO-COMBINATORIO* Extraído desde 24/10/2015.

117. Salamó, M. y Golobardes, E. (2004). Dynamic update of experience for a Case-Based Reasoning system. Extraído desde 9/016/2014, de http://www.maia.ub.edu/~maria/papers/Salamo2004a.pdf.

118. Sánchez, N. M., Lorenzo, M. M. G. y Valdivia, Z. Z. G. (2009). Modelo para diseñar sistemas de enseñanza-aprendizaje inteligentes utilizando el razonamiento basado en casos. *Revista Avances en Sistemas e Informática, 6*, 67-78.

119. Sarkar, A., Sahoo, G. y Sahoo, U. C. (2012). Application of fuzzy logic in transport planning. *3*, 1-21. Extraído desde 23/06/2015, de http://airccse.org/journal/ijsc/papers/3211ijsc01.pdf.

120. Schank, R. (1982). Dynamic Memory. *Cambridge Unviersity Press*. Extraído desde, de.

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

122. Shoniregun, C. A. (2006). Impacts and Risk Assessment of Technology for Internet Security: Enabled Information Small-Medium Enterprises (TEISMES) Vol. 17.(pp. 214). Extraído desde 20/10/2013, de http://books.google.com.cu/books?id=QkXpz2CtJp4C&pg=PA67&dq=risk+assessment+security+%2B+expert+system&hl=es&sa=X&ei=3g_rUp-KBYbyyAHi_YFo&ved=0CEwQ6AEwAw#v=onepage&q=risk%20assessment%20security%20%2B%20expert%20system&f=false.

123. Sikianakis, E. C., Antonakis, N. y Stolen, K. (2003). The CORAS approach for model-based risk management applied to a telemedicine service Vol. 95. Press, I. (Ed.) *The new navigators: from professional to patients: proceedings of MIE2003* (pp. 206). Extraído desde 12/12/2013, de http://www.google.com.cu/books?hl=es&lr=&id=i0KEzajbRUIC&oi=fnd&pg=PA206&dq=CORAS%2Brisk&ots=MUpRcseSo7&sig=bmgPKT6xDS_8-RxTLP6-WdHXsbE&redir_esc=y#v=onepage&q=CORAS%2Brisk&f=false.

124. Siler, W. y Buckley, J. J. (2005). *Fuzzy expert systems and fuzzy reasoning*. John Wiley & Sons, Inc. New Jersey, EUA.422

125. Smith, S. T. (1989). Modeling risk assessment applications with LAVA (Los Alamos Vulnerability/Risk Assessment)(pp. Medium: ED; Size: Pages: 5). Extraído desde 23/11/2013, de http://library.lanl.gov/cgi-bin/getfile?00222354.pdf.

126. Smiti, A. y Elouedi, Z. (2010). COID: Maintaining Case Method Based on Clustering, Outliers and Internal Detection. *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing 2010, 295*, 39–52.

127. Smiti, A. y Elouedi, Z. (2014). WCOID-DG: An approach for case base maintenance based on Weighting, Clustering, Outliers, Internal Detection and Dbsan-Gmeans. *Journal of Computer and System Sciences, 80*, 27-38.

128. Sridevi, B. y Nadarajan, R. (2009). Fuzzy Similarity Measure for Generalized Fuzzy Numbers. *Open Problems Compt. Math, 2*, 242-253. Extraído desde 10/06/2016, de http://www.emis.ams.org/journals/IJOPCM/Vol/09/IJOPCM(vol.2.2.7.J.9).pdf.

129. Stoneburner, Goguen, G. y Feringa, A. (2002). Risk management guide for information technology systems. *Nist special publication, 800*, 800-830. Extraído desde 7/12/2013, de www.security-science.com/pdf/risk-management-guide-for-information-technology-systems.pdf.

130. Talabis, M. y Martin, J. (2012). Information Security Risk Assessment Toolkit: Practical Assessments through Data Collection and Data Analysis Newnes (Ed.)(pp. 278). Extraído desde, de http://books.google.com.cu/books?id=0SctwiFj17AC&pg=PA49&dq=ISO+27005&hl=es&sa=X&ei=UE8AU4XSOOeCyAGa8IDABA&ved=0CDkQ6AEwAg#v=onepage&q=ISO%2027005&f=false.

131. Technologies, S. (2013). Extraído desde 18/01/2014, de http://www.softtreetech.com/dbaudit/index.htm.

132. Tóth, G. N. y Berek, L. (2010). Risk Methods and chance of practices. *Hungarian Journal of Industrial Chemistry, 38*, 193-196. Extraído desde 12/12/2013, de http://www.google.com.cu/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&ved=0CCkQFjAA&url=http%3A%2F%2Fkonyvtar.uni-pannon.hu%2Fhjic%2FHJIC38_193_196.pdf&ei=fZWfUq3hJpOMyAHAjYCwCQ&usg=AFQjCNHzTS0roSUWeYkyrl_wn3pcqkr2OQ&bvm=bv.57155469,d.aWc.

133. 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.

134. 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.

135. 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.

136. Wei, S.-H. y Chen, S.-M. (2009). A new approach for fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. *Expert Systems with Applications, 36*, 589 - 598. Extraído desde 14/06/2016, de http://www.sciencedirect.com/science/article/pii/S0957417407004630.

137. Whitman, M. y Mattord, H. (2011). Roadmap to Information Security: For IT and Infosec Managers: For IT and InfoSec Managers(pp. 400). Extraído desde 13/12/2013, de http://books.google.com.cu/books?id=UDE_T26hBm4C&pg=PA134&dq=:+Operationally+Critical+Threat,+Asset,+and+Vulnerability+Evaluation&hl=es&sa=X&ei=Zo2qUrOCKMv5kQeJrYGgCQ&ved=0CD0Q6AEwAg#v=onepage&q=%3A%20Operationally%20Critical%20Threat%2C%20Asset%2C%20and%20Vulnerability%20Evaluation&f=false.

138. Wong, L. R. P. y Mauricio, D. S. (2012). *Un modelo de razonamiento basado en casos para la captación de requisitos en el desarrollo de software.* Tesis de maestría, Universidad Nacional Mayor de San Marcos. Extraído desde 3/5/2014.

139. Wong, L. R. P., Mauricio, D. S. y Papa, E. A. (2014). Un modelo de Razonamiento Basado en Casos para la captación de requisitos en el desarrollo de proyectos de software. *Revista de investigación de Sistemas e Informática, 8*, 27-36. Extraído desde 27/11/2014, de http://revistasinvestigacion.unmsm.edu.pe/index.php/sistem/article/download/6323/5542.

140. Xu, Z., Shang, S., Qian, W. y Shu, W. (2010). A method or fuzzy risk analysis based on the new similarity of trapezoidal fuzzy numbers *Expert Systems with Applications, 37*, 1920-1927. Extraído desde 10/06/2016, de http://www.sciencedirect.com/science/article/pii/S0957417411001217.

141. Zhang, J., Lu, J. y Zhang, G. (2011). A Hybrid Knowledge-based Risk Prediction Method Using Fuzzy Logic and CBR for Avian Influenza Early Warning. *Journal of Multiple-Valued Logic & Soft Computing, 17*, 363-386. Extraído desde 16/07/2013, de http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=83c1714d-fa11-46ea-a0f3-c955d04587e0%40sessionmgr14&vid=1&hid=19.

142. Zhu, L.-S. y Xu, R.-N. (2012). Fuzzy risk analysis based on similarity measure of generalizard fuzzy numbers. Fuzzy Enginering and Operations Research *147*, 569-587. Extraído desde 10/06/2016, de http://link.springer.com/chapter/10.1007/978-3-642-28592-9_60.

ANEXOS

Anexo 1. Encuesta diagnóstico a especialistas en auditorías de seguridad de la información a sistemas computacionales.

Estimado auditor:

La presente encuesta responde a una investigación que se desarrolla en la Universidad de las Ciencias Informáticas, con el objetivo de caracterizar el proceso de evaluación del riesgo en las auditorías de seguridad informática a las Tecnologías de la Información (TI), llevado a cabo por los especialistas en auditorías de seguridad de la información a sistemas computacionales.

Solicitamos a Ud. que responda esta encuesta del modo más objetivo posible con el fin de obtener datos confiables que serán de suma importancia para lograr nuestros objetivos. Esta información tiene carácter anónimo y no se reportarán datos individuales.

1. **Datos Generales. Marque con una X según corresponda.**

01	¿Cuál es su nivel en cuanto a formación académica?	1. Técnico medio_____ 2. Universitario _____ 3. Master_____ 4. Doctor _____ 5. Otra _____ ¿Cuál?_____
02	Tiempo de experiencias como auditor en seguridad informática.	1. Menos de 5 años _____ 2. De 5 a 10 años_____ 3. Más de 10 años _____
03	La formación como auditor en seguridad informática es:	1. Empírica_____ 2. Por cursos _____ 3. Diplomados _____ 4. Otra _____ ¿Cuál? _____

2. Proceso de evaluación del riesgo de seguridad de la información en las TI. Marque con una X según corresponda.

04	¿Cómo realiza su organización la detección de las vulnerabilidades?	1. Aplicación de encuestas ____ 2. Monitoreo a las TI ____ 3. Otra____
05	¿Se utiliza alguna herramienta para el monitoreo de las TI?	1.Encuestas____ 2. Listas de chequeo____ 3. Microsoft Baseline Security Analyzer ____ 5. MSAT ____ 6. Microsoft Forefront____ 7. DB Audit ____ 8. Secure Oracle Auditor ____ 9. Secure SQL Auditor ____ 10. Otra____ ¿Cuál? _____
06	¿Qué metodología, técnica o estándar se utilizan para la evaluación del riesgo de seguridad de la información?	1. Ninguna ____ 2. COBRA ____ 3. CORAS ____ 4. EBIOS ____ 5. BPIRM ____ 6. IT-Grundschutz (IT Baseline Protection Manual) ____ 7. CRAMM ____ 8. ISRAM ____ 9. NIST SP 800-53____ 10. ISAMM ____ 11. LAVA____ 12. OCTAVE ____

		13. MAGERIT ____ 14. SOMAP ____ 15. ISO/IEC 27001/27002 ____ 16. Cobit ____ 16. Otra ____ ¿Cuál? _____
07	¿Qué software utiliza para la evaluación de riesgo?	1. Ninguno ____ 2. Callio Secura ____ 3. CASIS ____ 4. Control Compliance Suite (CCS) 11 ____ 5. Citicus ONE ____ 6. RiskWatch 360____ 7. COBRA ____ 8. CRAMM ____ 9. ISRAC ____ 10. Tuar ____ 11. RAMEX____ 12. Sistema AGRIS ____ 13. (CIS-CAT) ____ 14. CobitAdvisore____ 15. Otra ____ ¿Cuál? _____
08	El resultado de la evaluación de riesgo se le proporciona a la entidad auditada en:	1. Palabras ____ 2. Números ____ 3. Ninguna ____
09	¿Actualmente cuánto se tarda en llegar al resultado de la evaluación del riesgo de la TI contando desde el inicio de la auditoría?	1. Minutos ____ 2. Horas____ 3. Días____
10	¿Con qué frecuencia se recurre a la experiencia en las auditorías de seguridad	1. Siempre ____

	informáticas en la evaluación del riesgo?	2. Muchas o casi siempre ____ 3. Pocas veces ____ 4. Ninguna ____
11	Sobre el posible nivel de subjetividad de las auditorías: ¿Ud. considera que puede existir diferencia en la evaluación del riesgo de seguridad de la información en dependencia de los especialistas que participan?	1. Ninguna____ 2. Pocas____ 3. Llega a cambiar el resultado de la evaluación ____

3. **Recomendaciones generales. Marque con una X según corresponda.**

12	¿Cree usted que sería útil un software que automatice la evaluación del riesgo de seguridad de la información en las TI?	1. Sí ____ 2. No ____
13	¿Considera que sería más efectivo el proceso de evaluación del riesgo de seguridad de la información en las TI, si se pudiera contar en el momento de la auditoría con experiencias pasadas de apoyo a la toma de decisiones?	1. Sí ____ 2. No ____
14	¿Cree que es importante la utilización de fórmulas matemáticas para determinar el riesgo de seguridad de la información en las TI?	1. Sí ____ 2. No ____

4. **Expliquebrevemente el procedimiento, con que variables y valores se evalúanlos parámetros de la lista de chequeo:**

Sobre la aplicación de la encuesta:

La encuesta fue aplicada a especialistas en auditorías de seguridad de la información a sistemas computacionalesen Cuba y con el objetivo fundamental de caracterizar el proceso de evaluación del riesgo en las auditorías de seguridad informática a las Tecnologías de la Información (TI).

Anexo 1A. Entrevista diagnóstico a especialistas en auditorías de seguridad de la información a sistemas computacionales

Objetivo: Evaluar el estado del proceso de evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos.

Guía de observación:

1. ¿Existe uno o variosprocesospara la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos? ¿Cuál utilizan?

2. ¿Este proceso se ejecuta con la misma calidad ose obtienen los mismos resultados con todos los auditores conocidos?¿En caso de existir diferencias pudiera identificar cuáles son las causas?

3. ¿Conoce alguna solución que permita realizar la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos con resultados semejantes a las de un experto?

4. ¿Utilizan alguna herramienta informática, técnica o metodología que apoye la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos? ¿En caso de identificar alguna, pudiera cual es la ventaja y desventaja de la misma?

5. ¿Considera usted necesario el perfeccionamiento del proceso parala evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos a partir de sistemas automatizados?

6. ¿Considera usted conveniente incorporar la experiencia del auditorpara aumentar la exactitud y disminuir el tiempo de respuesta del proceso de evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos?

Anexo 2. Composición de expertos involucrados en el diagnóstico.

Número	Perfil de expertos	Cantidad
1	**Años de experiencia en su actividad**	
	Menos de 5 años	10
	De 5 a 10 años	4
	Más de 10 años	5
2	**Nivel en cuanto a formación académica**	
	Técnico medio	3
	Universitario	16
	Master	0
	Doctor	0
	Otra	0
3	**Formación como auditor en seguridad informática**	
	Empírica	16
	Por cursos	4
	Diplomados	0
	Otra	0

Anexo 3.Rasgosde la lista de chequeo para el SGBD Microsoft SQL Server 2000.

Rasgos	Impacto	Peso	Evaluación	Riesgo
1.1: Cuentas con seguridad integrada.	Alto	0.69	Bien, Regular o Mal	Alto, Medio o Bajo
1.2: Longines (ingresos)del motor de la base de datos.	Alto	0.75		
1.3: Cuentas genéricas dentro del motor de base de datos.	Medio	0.6		
1.4: Usuarios con claves nulas o triviales.	Alto	1		
1.5: Verificar si existe en el motor de base de datos el Grupo BUILTIN \POWER USERS.	Medio	0.66		
1.6: Roles del motor de base de datos.	Alto	0.7		
1.7: Segregación de funciones.	Alto	0.68		
1.8: Modificación del esquema de bases de datos.	Alto	1		
1.9: Permisos sobre tablas sensitivas del motor de BD.	Alto	1		
2.1: Utilización de controladores de dominio.	Medio	0.4		
2.2: Permisos sobre directorios de la base de datos.	Alto	1		
2.3: Ejecución de otras aplicaciones.	Medio	0.37		
2.4: Utilitarios para el manejo y administración de la base de datos.	Bajo	0.04		
2.5: Bases de datos genéricas.	Medio	0.4		
2.6: Salvas periódicas y bases de datos en desuso	Alto	0.7		
2.7: Conexión con otros servidores.	Medio	0.37		
2.8: Cuentas de Servicio.	Medio	0.37		

2.9: Configuración del SQL Server Agent.	Medio	0.66		
2.10: Verificar la versión del motor de bases de datos y si están instalados los servicepack.	Alto	0.7		
2.11: Modificaciones directa sobre tablas del sistema.	Alto	1		
2.12: Acceso a instrucciones y objetos.	Alto	1		
2.13: Acceso remoto a procedimientos almacenados.	Alto	1		
2.14: Ejecución automática de procedimientos.	Alto	0.7		
2.15: Auditoría.	Alto	1		
3.1: Permisos de usuarios finales.	Alto	1		
3.2: Base de datos por default.	Alto	1		
3.3: Roles en la base de datos.	Medio	0.6		
3.4: Accesos de personal de desarrollo a bases productivas.	Alto	1		
3.5: Cuenta de usuarios Guest.	Alto	1		
3.6: Backup de las bases de datos.	Alto	0.68		
3.7: Servicio de transformación de datos (DTS).	Medio	0.64		

Anexo4.Encuesta para determinar los pesos de los parámetros y las escalas de los valores lingüísticos.

Estimado auditor:

La presente encuesta responde a una investigación que se desarrolla en la Universidad de las Ciencias Informáticas, con el objetivo de identificar las diferencias en cuanto al impacto entre los parámetros que componen las listas de chequeo de seguridad de los Sistemas Gestores de Bases de Datos y de los valores lingüísticos de esta variable y del riesgo. Estos datos son importantes enel proceso de evaluación del riesgo en las auditorías de seguridad informática a las Tecnologías de la Información (TI), llevado a cabo por los especialistas en auditorías de seguridad de la información a sistemas computacionales.

Solicitamos a Ud. que responda esta encuesta del modo más objetivo posible con el fin de obtener datos confiables que serán de suma importancia para lograr el objetivo trazado.

1. El procedimiento consiste en señalar con una x en la escala el lugar en el cual usted ubicaría su opinión acerca de los valores que le corresponden de los siguientes parámetros.

Parámetro	**Impacto**	**Marque con una X**
1.1:Cuentas con seguridad integrada	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algoalto _____
1.2: Longines (ingresos) del motor de la base de datos	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
1.3: Cuentas genéricas dentro del motor de base de datos	¿Qué tan medio es el	Medio muy cerca del Alto _____ Medio cerca del Alto _____ Medio_____

	impacto?	Medio cerca del bajo _____ Medio muy cerca del bajo _____
1.4: Usuarios con claves nulas o triviales	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
1.5:Verificar si existe en el motor de base de datos el Grupo BUILTIN \POWER USERS	¿Qué tan medio es el impacto?	Medio muy cerca del Alto _____ Medio cerca del Alto _____ Medio_____ Medio cerca del bajo _____ Medio muy cerca del bajo _____
1.6: Roles del motor de base de datos.	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Más o menos alto _____ Algo alto _____
1.7: Segregación de funciones	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
1.8: Modificación del esquema de bases de datos	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____

		Algo alto _____
1.9: Permisos sobre tablas sensitivas del motor de BD	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
2.1: Utilización del PrimaryDomainController	¿Qué tan medio es el impacto?	Medio muy cerca del Alto _____ Medio cerca del Alto _____ Medio_____ Medio cerca del bajo _____ Medio muy cerca del bajo _____
2.2: Permisos sobre directorios de la base de datos.	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
2.3: Ejecución de otras aplicaciones.	¿Qué tan medio es el impacto?	Medio muy cerca del Alto _____ Medio cerca del Alto _____ Medio_____ Medio cerca del bajo _____ Medio muy cerca del bajo _____
2.4: Utilitarios para el manejo y administración de la base de datos	¿Qué tan bajo es el impacto?	Extremadamente bajo _____ Muy bajo _____ Bajo_____ Más o menos bajo _____

		Algo bajo _____
2.5: Bases de datos genéricas	¿Qué tan medio es el impacto?	Medio muy cerca del Alto _____ Medio cerca del Alto _____ Medio_____ Medio cerca del bajo _____ Medio muy cerca del bajo _____
2.6 Salvas periódicas y bases de datos en desuso.	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
2.7 Conexión con otros servidores	¿Qué tan medio es el impacto?	Medio muy cerca del Alto _____ Medio cerca del Alto _____ Medio_____ Medio cerca del bajo _____ Medio muy cerca del bajo _____
2.8 Cuentas de servicio	¿Qué tan medio es el impacto?	Medio muy cerca del Alto _____ Medio cerca del Alto _____ Medio_____ Medio cerca del bajo _____ Medio muy cerca del bajo _____
2.9: Configuración del SQL Server Agent	¿Qué tan medio es el impacto?	Medio muy cerca del Alto _____ Medio cerca del Alto _____ Medio_____ Medio cerca del bajo _____

		Medio muy cerca del bajo _____
2.10: Verificar la versión del motor de bases de datos y si están instalados los service pack	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
2.11: Modificaciones directa sobre tablas del sistema	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
2.12: Acceso a instrucciones y objetos	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
2.13: Acceso remoto a procedimientos almacenados	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
2.14: Ejecución automática de procedimientos	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____

		Algo alto _____
2.15: Auditoría	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
3.1: Permisos de usuarios finales	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
3.2: Base de datos por default	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
3.3: Roles en la base de datos	¿Qué tan medio es el impacto?	Medio muy cerca del Alto _____ Medio cerca del Alto _____ Medio_____ Medio cerca del bajo _____ Medio muy cerca del bajo _____
3.4: Accesos de personal de desarrollo a bases productivas	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____

		Algo alto _____
3.5 Cuenta de usuarios Guest	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
3.6: Backup de las bases de datos	¿Qué tan alto es el impacto?	Extremadamente alto _____ Muy alto _____ Alto_____ Más o menos alto _____ Algo alto _____
3.7 Servicio de transformación de datos (DTS)	¿Qué tan medio es el impacto?	Medio muy cerca del Alto _____ Medio cerca del Alto _____ Medio_____ Medio cerca del bajo _____ Medio muy cerca del bajo _____

2. Especifique en una escala del 0 a 1 en la cual usted considera que debe estar presente los siguientes valores lingüísticos:

a. Alto: valor mínimo_____, valor máximo _____

b. Medio: valor mínimo_____, valor máximo _____

c. Bajo:valor mínimo_____, valor máximo _____

Escala para el alto:

- Extremadamente alto ___1__
- Muy alto __0.75___

- Alto___0.7__
- Más o menos alto __0.69___
- Algo alto __0.68___

Escala para el Medio

- Medio muy cerca del Alto __0.66___
- Mediocerca del Alto__0.64___
- Medio__0.6___
- Medio cerca delbajo __0.4___
- Medio muy cerca delbajo __0.37___

Escala para el Bajo

- Extremadamente bajo __0___
- Muy bajo ___0.04__
- Bajo___0.25__
- Más o menos bajo __0.28___
- Algo bajo __0.3___

Anexo5.Encuesta para determinar reglas difusas de evaluación.

Encuesta para esclarecer la forma de proceder en la evaluación de la auditoría de seguridad informática en los servidores de bases de datos.

A la hora de auditar la base de datos de una aplicación se encuentra que está hospedada en más de un servidor. Evidentemente hay que auditar a todos los servidores donde se encuentrala base de datos. Pero una vez obtenida la evaluación de cada servidor o matriz de resultado.¿Cómo se determina el resultado final en el informe de la auditoría de bases de datos teniendo en cuenta que es más de un servidor lo que fue auditado? ¿Cómo se evalúa si son don o 3 o más servidores con evaluaciones diferentes?

Se ponen en forma de reglas, varios ejemplo para ver si es alguna de estas. Si es de otra manera, escríbela con el mismo estilo. Los ejemplos son:
Si hay tres servidores evaluados:

1. Si todos los servidores son evaluados de Medio entonces la evaluación en el informe es de Medio.
2. O si al menos un servidor es evaluado de Alto entonces la evaluación en el informe es de Alto.
3. O si el **promedio** de la evaluación de los servidores es de Bajo entonces la evaluación en el informe es de Bajo.

¿Usted cree que utilizar el promedio de las evaluaciones, como resultado final es una buena solución o no?

Resultados de la encuesta

- Se decide por la segunda variante de las reglas presentadas.

- El promedio no sirve para evaluar, pues puede dar un servidor bajo y otro alto y la evaluación entonces sería medio, cuando en realidad por ese servidor que dio alto se puede vulnerar la aplicación.

- Tomar siempre la evaluación más crítica y asumir esta como la evaluación final. Un parámetro de la lista de chequeo con evaluación alta, pone en un estado de riesgoalto el servidor de base de datos auditado.

Anexo 6.Cuestionario aplicado para el método de consultas a experto.

Nombre y apellidos del especialista_____

Sector al que pertenece: Académico: ___ Empresarial: ___ Gubernamental: ___

Nivel educacional: técnico medio_____, universitario _____, máster_____, doctor _____, otra _____, ¿cuál?_____

Cantidad de años de experiencia en la seguridad informática: ____

Certificaciones obtenidas relacionadas con seguridad informática:_____

Cantidad de publicaciones científicas relacionadas con la seguridad informática en los últimos 7 años: ____

Cantidad de eventos científicos en los que ha participado en los últimos 7 años: ____

En el documento adjunto se presenta un resumen del modelo propuesto para la evaluación del riesgo de seguridad de la información

Sobre esta propuesta evalúe el modelo con respecto a las siguientes preguntas(asigne valores del 1 al 5, donde De acuerdode forma Excelente-5, Muy de acuerdo-4, De acuerdo-3, Indeciso-2, En desacuerdo-1):

Preguntas	**Evaluación**
1. ¿Considera que el primer componente satisface la fase del monitoreo?	
2. ¿Considera adecuada la forma de determinar la evaluación de los parámetros de la lista de chequeo del segundo componente?	
3. ¿Manifieste su opinión sobre la propuesta del cálculo del riesgo?	
4. ¿Aprecia que puede resultar favorable la introducción de aspectos de la lógica difusa y el razonamiento basado en caso (tal como se propone en el modelo) para lograr una mayor exactitud y disminución del tiempo de respuesta en la evaluación del RSI en los SGBD?	
5. ¿El modelo en su conjunto es capaz de evaluar el riesgo de seguridad de la información en los sistemas gestores de bases de datos?	
6. ¿Considera adecuada la utilización de fórmulas y de algoritmos basados en el reconocimiento de patrones para agilizar y calibrar automáticamente el modelo?	

7. ¿Considera pertinente la estructura del modelo?	
8. ¿Considera necesario incorporar algún otro componente al modelo?	

9. ¿Considera usted que se deba evaluar el riesgo de seguridad de la información en los sistemas gestores de bases de datos sin un modelo que oriente el proceso?

   Sí_____ b) No sé _____ c) No_____

10. ¿Si usted necesitara realizar la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos usaría este modelo?

    Sí_____ b) No sé _____ c) No_____

11. ¿Le satisface este modelo para la evaluación del riesgo de seguridad de la información en los sistemas?

    a) Me gusta mucho _____ b) No me gusta tanto _____

    c) Me da lo mismo _____ d) Me disgusta más de lo que me gusta _____

    e) No me gusta nada _____ f) No sé qué decir _____

12. ¿Considera que con la aplicación del modelo aumentaría la exactitud en la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos?

    __ Mucho __ Bastante __ Algo __ Casi nada __ Nada

13. ¿Considera que con la aplicación del modelo se disminuye el tiempo de respuesta en la evaluación del riesgo de seguridad de la información en los sistemas gestores de bases de datos?

    __ Mucho __ Bastante __ Algo __ Casi nada __ Nada

14. ¿Desea añadir algún otro comentario sobre el modelo propuesto?

Anexo 7. Determinación del nivel de competencia de los expertos.

La competencia de los expertos se determina por el coeficiente K, el cual se calcula de acuerdo con la opinión del candidato sobre su nivel de conocimiento acerca del problema que se está resolviendo y con las fuentes que le permiten argumentar sus criterios.

El cálculo del coeficiente de concordancia de Kendall se realiza de la forma siguiente:

K = ½ (Kc + Ka)

Donde:

K:Coeficiente de competencia.

Kc: Es el coeficiente de conocimiento o información que tiene el experto acerca del problema, el cual es calculado sobre la base de la valoración del propio experto en una escala de 0 a 10 y multiplicado por 0.1.

- El valor 0 indica absoluto desconocimiento de la problemática que se evalúa.
- El valor 1 indica pleno conocimiento de la referida problemática

0	1	2	3	4	5	6	7	8	9	10

El experto deberá marcar una cruz en la casilla que estime pertinente según suautovaloración del nivel de conocimiento que posee en el tema tratado.
Ka: Es el coeficiente de argumentación o fundamentación de los criterios del experto, determinado como resultado de la suma de los puntos alcanzados a partir de una tabla patrón.
La tabla patrón forma parte del cuestionario que se le aplica a los candidatos a expertos y en la misma estos reflejan el grado de influencia de las fuentes mediante las cuales han asimilado los conocimientos sobre el tema objeto de valoración.

Fuentes de argumentación		Grado de influencia de cada una de las fuentes en sus criterios		
		Alto	**Medio**	**Bajo**
1.1	Preparación en seguridad informática para las Tecnologías de la Información.	0.3	0.2	0.1
1.2	Experiencia adquirida durante su vida profesional en seguridad informática en las Tecnologías de la Información.	0.5	0.4	0.2
1.3	Nivel de conocimiento teórico en seguridad informática en las Tecnologías de la Información.	0.05	0.05	0.05
1.4	Actualización en cursos de postgrado, diplomados, maestrías, doctorado, etc.	0.05	0.05	0.05
1.5	Conocimiento propio sobre la seguridad informática en las Tecnologías de la Información.	0.05	0.05	0.05
1.6	Intuición.	0.05	0.05	0.05

Al experto se le presenta esta tabla sin cifras, orientándole que marque con una (x) sobre las fuentes que han influido más en su conocimiento de acuerdo con los niveles ALTO (A), MEDIO (M) y BAJO (B). Posteriormente utilizando los valores de la tabla patrón para cada una de las celdas marcadas por el experto, se calcula el número de puntos obtenidos en total.
Finalmente se determina el coeficiente de competencia K de modo que:

Si 0,8< K < 1 el experto tiene competencia alta.
Si 0,5 < K < 0,8 el experto tiene competencia media.
Si 0 < K <0,5 el experto tiene competencia baja.

RESULTADOS: Competencia de los 17expertos que participaron en la validación del modelo para la evaluación del RSI en los SGBD.

Expertos	Kc	1.1	1.2	1.3	1.4	1.5	1.6	Ka	K	CC
1	1	0.3	0.5	0.05	0.05	0.05	0.05	1	1	Alto
2	0.4	0.1	0.2	0.05	0.05	0.05	0.05	0.9	0.65	Medio
3	0.6	0.1	0.4	0.05	0.05	0.05	0.05	0.7	0.65	Medio
4	1	0.3	0.5	0.05	0.05	0.05	0.05	1	1	Alto
5	1	0.3	0.5	0.05	0.05	0.05	0.05	1	1	Alto
6	0.4	0.1	0.2	0.05	0.05	0.05	0.05	0.9	0.65	Medio
7	0.7	0.1	0.2	0.05	0.05	0.05	0.05	0.5	0.6	Medio
8	1	0.3	0.5	0.05	0.05	0.05	0.05	1	1	Alto
9	0.5	0.1	0.2	0.05	0.05	0.05	0.05	0.5	0.5	Bajo
10	0.8	0.1	0.4	0.05	0.05	0.05	0.05	0.7	0.75	Medio
11	0.8	0.1	0.4	0.05	0.05	0.05	0.05	0.7	0.75	Medio
12	0.9	0.1	0.5	0.05	0.05	0.05	0.05	0.8	0.85	Alto
13	1	0.3	0.5	0.05	0.05	0.05	0.05	1	1	Alto
14	0.8	0.1	0.4	0.05	0.05	0.05	0.05	0.7	0.75	Medio
15	0.8	0.3	0.4	0.05	0.05	0.05	0.05	0.9	0.85	Alto
16	1	0.3	0.5	0.05	0.05	0.05	0.05	1	1	Alto
17	0.8	0.1	0.4	0.05	0.05	0.05	0.05	0.7	0.75	Medio

Alto	8
Medio	8
Bajo	1
Total	17

Anexo 8. Listado de los especialistas que participaron en la consulta de expertos.

Núm.	**Experto**	**Entidad**	**Cargo**	**Experiencia (años)**
1	**Nancy Vandama**	ETECSA	Jefe de Departamento	20
2	Olivia Rodríguez Abril	ETECSA	Auditor en seguridad informática	10
3	Viviana Pozo Contreras	ETECSA	Auditor en seguridad informática	12
4	**Alexander López Gavilán**	ETECSA	Auditor en seguridad informática	14
5	**Liera González**	ETECSA	Auditor en seguridad informática	12
6	Ángel EduardoPentín S.	ETECSA	Auditor en seguridad informática	2
7	Ángel Miguel Fuego Calvo	ETECSA	Auditor en seguridad informática	1
8	**Blas A. Tejeda Mendoza**	ETECSA	Auditor en seguridad informática	3
9	Camila Argüelles Alonso	ETECSA	Auditor en seguridad informática	1
10	**Carlos Gabriel PrimellesEguía**	ETECSA	Auditor en seguridad informática	1
11	**Cecilia Oro Rivera**	ETECSA	Auditor en seguridad informática	6
12	**Gerardo Broughton Sáez**	ETECSA	Auditor en seguridad informática	12
13	**IntiJimenez**	ETECSA	Auditor en seguridad informática	7
14	**JenyFerranro**	ETECSA	Auditor en seguridad informática	7
15	**José A. León Martínez**	ETECSA	Auditor en seguridad informática	4
16	**LeonardoLaugatChibás**	ETECSA	Auditor en seguridad informática	4
17	MarianelaOzco Pérez	ETECSA	Auditor en seguridad informática	2

Anexo 9. Resultados de la aplicación de la encuesta a expertos.

Aspectos básicos evaluados en el modelo		**Cantidad Excelente + muy de acuerdo**	**Cantidad en %**
1	Estructura del modelo	10	91
2	La evaluación de los parámetros	8	73
3	El cálculo del riesgo	8	73
4	El uso de las técnicas de inteligencia artificial	10	91
5	La utilización de fórmulas y de algoritmos basados en el reconocimiento de patrones	9	82
6	Modelo en su conjunto	10	91

Anexo 10.Diseño de los casos de estudio para la validación del modelo.

Parámetro	**O_5**	**O_6**	**O_7**	**O_8**	**O_9**	**O_{10}**	**O_{11}**	**O_{12}**	**O_{13}**	**O_{14}**	**O_{15}**
1.1:	Alto	Alto	Alto	Alto	Bajo	Bajo	Bajo	Alto	Alto	Bajo	Alto
1.2:	Bajo	Bajo	Bajo	Alto	Bajo	Bajo	Bajo	Alto	Alto	Alto	Alto
1.3:	Bajo	Bajo	Bajo	Bajo	Medio	Medio	Bajo	Medio	Medio	Medio	Medio
1.4:	Bajo	Bajo	Bajo	Alto	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Alto
1.5:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo
1.6:	Alto	Alto	Alto	Alto	Bajo	Bajo	Bajo	Alto	Alto	Alto	Alto
1.7:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo
1.8:	Alto	Bajo	Bajo	Bajo	Bajo	Bajo	Alto	Alto	Alto	Alto	Alto
1.9:	Alto	Alto	Alto	Alto	Bajo	Bajo	Bajo	Bajo	Alto	Bajo	Bajo
2.1:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo
2.2:	Bajo	Bajo	Alto	Alto	Bajo	Bajo	Bajo	Bajo	Alto	Alto	Alto
2.3:	Medio	Medio	Medio	Medio	Medio	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo
2.4:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo
2.5:	Bajo	Bajo	Bajo	Bajo	Medio	Bajo	Medio	Bajo	Bajo	Bajo	Alto
2.6:	Alto	Alto	Bajo	Alto	Alto	Alto	Bajo	Bajo	Bajo	Bajo	Bajo
2.7:	Bajo	Bajo	Bajo	Bajo	Medio	Bajo	Bajo	Bajo	Alto	Bajo	Bajo
2.8:	Bajo	Bajo	Bajo	Bajo	Medio	Bajo	Bajo	Medio	Medio	Bajo	Alto
2.9:	Bajo	Bajo	Bajo	Bajo	Medio	Bajo	Bajo	Bajo	Bajo	Bajo	Alto
2.10:	Alto	Alto	Alto	Bajo	Bajo	Bajo	Bajo	Alto	Bajo	Bajo	Bajo
2.11:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo
2.12:	Alto	Alto	Alto	Alto	Bajo	Bajo	Bajo	Bajo	Alto	Alto	Bajo
2.13:	Alto	Alto	Alto	Alto	Bajo	Bajo	Bajo	Bajo	Alto	Alto	Bajo
2.14:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Alto	Bajo	Bajo
2.15:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Alto	Bajo	Bajo	Bajo
3.1:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Alto	Bajo	Bajo	Alto
3.2:	Alto	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Alto	Alto	Bajo	Alto
3.3:	Bajo	Bajo	Medio	Medio	Bajo	Bajo	Bajo	Bajo	Bajo	Medio	Bajo
3.4:	Alto	Alto	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Alto
3.5:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Alto	Alto	Bajo	Bajo
3.6:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Alto	Bajo	Bajo	Bajo	Alto
3.7:	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Bajo	Medio	Alto	Bajo	Bajo
Evaluación esperada	**Alto**	**Alto**	**Alto**	**Alto**	**Bajo**	**Bajo**	**Bajo**	**Alto**	**Alto**	**Alto**	**Alto**

Anexo 11. Resultado de las pruebas de normalidad con el SPSS v13.0.

Hipótesis estadísticaspara la propuesta de Patra y Mondal:

- **Hipótesis nula H_o**: el conjunto de datos sigue una distribución normal para la propuesta de Patra y Mondal.
- **Hipótesis alternativa H1:** el conjunto de datos no sigue una distribución normal para la propuesta de Patra y Mondal.

Pruebas de normalidad							
Calificación		**Kolmogorov-Smirnov(a)**			**Shapiro-Wilk**		
		Estadístico	**gl**	**Sig.**	**Estadístico**	**gl**	**Sig.**
Propuesta de Patra y Mondal	**Alto**	,314	9	,011	,771	9	,010
	Bajo	,187	9	,200(*)	,951	9	,702
	Medio	,241	9	,142	,870	9	,123
* Este es un límite inferior de la significación verdadera.							
a Corrección de la significación de Lilliefors							

Tabla A1. Pruebas de normalidad para los datos de la propuesta de Patra y Mondal.

Hipótesis estadísticaspara la propuesta del autor:

- **Hipótesis nula H_o**: el conjunto de datos sigue una distribución normal para la propuesta del autor.
- **Hipótesis alternativa H1:** el conjunto de datos no sigue una distribución normal para la propuesta del autor.

Pruebas de normalidad							
Calificación		**Kolmogorov-Smirnov(a)**			**Shapiro-Wilk**		
		Estadístico	**gl**	**Sig.**	**Estadístico**	**gl**	**Sig.**
Propuesta del autor	**Alto**	,305	9	,016	,730	9	,003
	Bajo	,186	9	,200(*)	,963	9	,828
	Medio	,204	9	,200(*)	,885	9	,177
* Este es un límite inferior de la significación verdadera.							
a Corrección de la significación de Lilliefors							

Tabla A2. Pruebas de normalidad para los datos de la propuesta del autor.

Anexo 12. Resultado de las pruebas estadísticas con el SPSS v13.0.

Prueba Tpara los datos con distribución normal.

Hipótesis estadísticas para la variable Alto:

- **Hipótesis nula H_0**: No existe una diferencia significativa en el conjunto de los valores de semejanzaentre la propuesta del autor y la de Patra y Mondal para la variable Alto.
- **Hipótesis alternativa H1:** Sí existe una diferencia significativa en el conjunto de los valores de semejanza entre la propuesta del autor y la de Patra y Mondal para la variable Alto.

Estadísticos de muestras relacionadas

		Media	N	Desviación típ.	Error típ. de la media
Par 1	**PM**	,31078	9	,222923	,074308
	Autor	,27678	9	,290564	,096855

Correlaciones de muestras relacionadas

		N	Correlación	Sig.
Par 1	**PM y Autor**	9	,995	,000

Prueba de muestras relacionadas

		Diferencias relacionadas					t	gl	Sig. (bilateral)
		Media	Desviación típ.	Error típ. de la media	95% Intervalo de confianza para la diferencia				
					Inferior	Superior			
Par 1	**PM - Autor**	,034000	,072069	,024023	-,021397	,089397	1,415	8	,195

Análisis de los datos: La media de la propuesta de Patra y Mondal es mayor que la del autor.Con la tabla de correlaciones se comprueba que la correlación existente entre las puntuaciones obtenidas en los dos parciales es de 0,995 y que la prueba de inferencia asociada es significativa, todo ello implica que los grupos eran efectivamente relacionados.La tabla de las muestras relacionadas indicaque no hay diferencias estadísticamente significativas entre la propuesta de Patra y Mondal y la del autor al mostrar un valor de confiabilidad > 0.05.

Prueba los rangos con signo de Wilcoxonpara los datos con distribución no normal.

Hipótesis estadísticas para la variable Medio:

- **Hipótesis nula H_o**: No existe una diferencia significativa en el conjunto de los valores de semejanzaentre la propuesta del autor y la de Patra y Mondal para la variable Medio.
- **Hipótesis alternativa H1:** Sí existe una diferencia significativa en el conjunto de los valores de semejanza entre la propuesta del autor y la de Patra y Mondal para la variable Medio.

Estadísticos descriptivos

	N	Media	Desviación típica	Mínimo	Máximo	Percentiles		
						25	50 (Mediana)	75
PM	9	,66978	,176939	,392	,847	,49150	,75500	,83550
Autor	9	,60289	,222737	,320	,883	,36300	,68800	,81700

Prueba de los rangos con signo de Wilcoxon

Rangos

		N	Rango promedio	Suma de rangos
Autor - PM	**Rangos negativos**	8(a)	5,25	42,00
	Rangos positivos	1(b)	3,00	3,00
	Empates	0(c)		
	Total	9		

a Autor < PM

b Autor > PM

c Autor = PM

Estadísticos de contraste(b)

	Autor - PM
Z	-2,310(a)
Sig. asintót. (bilateral)	,021

a Basado en los rangos positivos.
b Prueba de los rangos con signo de Wilcoxon

Análisis de los datos: La media de la propuesta de Patra y Mondal es mayor que la del autor. La tabla de las muestras relacionadas indica que hay diferencias estadísticamente significativas entre la propuesta de Patra y Mondal y la del autor al mostrar un valor de confiabilidad menor que 0.05. Por lo que se rechaza la hipótesis nula.

Hipótesis estadísticas para la variable Bajo:

- **Hipótesis nula H_o**: No existe una diferencia significativa en el conjunto de los valores de semejanzaentre la propuesta del autor y la de Patra y Mondal para la variable Bajo.
- **Hipótesis alternativa H1:** Sí existe una diferencia significativa en el conjunto de los valores de semejanza entre la propuesta del autor y la de Patra y Mondal para la variable Bajo.

Estadísticos descriptivos

	N	Media	Desviación típica	Mínimo	Máximo	Percentiles		
						25	50 (Mediana)	75
PM	9	,48300	,232175	,065	,844	,37300	,45200	,68850
Autor	9	,53511	,289876	,028	,964	,36900	,51700	,80600

Prueba de los rangos con signo de Wilcoxon

Rangos

		N	Rango promedio	Suma de rangos
Autor - PM	**Rangos negativos**	2(a)	2,75	5,50
	Rangos positivos	7(b)	5,64	39,50
	Empates	0(c)		
	Total	9		

a Autor < PM
b Autor > PM
c Autor = PM

Estadísticos de contraste(b)	
	Autor - PM
Z	-2,016(a)
Sig. asintót. (bilateral)	,044
a Basado en los rangos negativos.	
b Prueba de los rangos con signo de Wilcoxon	

Análisis de los datos: La media de la propuesta de Patra y Mondal es menor que la del autor. La tabla de las muestras relacionadas indica que hay diferencias estadísticamente significativas entre la propuesta de Patra y Mondal y la del autor al mostrar un valor de confiabilidad menor que 0.05. Por lo que se rechaza la hipótesis nula.