



Universidad de las Ciencias Informáticas
Facultad 2

Trabajo de Diploma para optar por el Título de
Ingeniero en Ciencias Informáticas

**PERSONALIZACIÓN DEL PLUGIN FUZZY MINER DE PROM
PARA DETECTAR RUIDO EN LOS PROCESOS DEL SISTEMA
DE INFORMACIÓN HOSPITALARIA DEL CESIM**

Autores: Katherine Pereiras Viera
Adrián Sosa Benítez

Tutor: Ing. Arturo Orellana García

La Habana, junio del 2015

“Año 57 de la Revolución”

DECLARACIÓN DE AUTORÍA

Declaramos ser autores de la presente tesis y reconocemos a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Para que así conste firmo la presente a los ____ días del mes de _____ del año _____.

Katherine Pereiras Viera

Autor

Adrián Sosa Benítez

Autor

Ing. Arturo Orellana García

Tutor



Si salgo llevo, si llevo entro, si entro triunfo.

Fidel Castro

Katherine

A mis padres, Belkys y Lázaro, a quienes con este logro quiero devolver un poco de lo que me han dado.

A mi familia, amigos y personas que incondicionalmente están siempre conmigo.

Adrián

A mi madre Marilyn.

A mi abuelo Enrique por ser mi ejemplo y guía.

A todas las personas de las que recibí un consejo y un apoyo para cada camino que emprendía.

RESUMEN

Las Tecnologías de Información y las Comunicaciones han permitido que las organizaciones utilicen sistemas automatizados para gestionar sus procesos. Estos almacenan acciones llamadas trazas referentes a las actividades que se gestionan. Los registros de eventos son el producto de extraer y transformar las trazas de ejecución del sistema. Con la Minería de Procesos se pueden analizar los registros de eventos a partir de la aplicación de técnicas existentes en forma de *plugins*¹, ejemplo *Fuzzy Miner*, que genera un modelo basado en grafos, capaz de detectar la presencia de ruido en un registro de eventos. Una interpretación de la presencia de ruido es que ocurran actividades anómalas en el proceso que atenten contra el correcto funcionamiento del sistema. La investigación se centró en la personalización del *plugin Fuzzy Miner* para detectar ruido en los procesos del Sistema de Información Hospitalaria del Centro de Informática Médica. Para su desarrollo se empleó Eclipse 3.4.2 como ambiente de desarrollo, como lenguaje de programación Java 1.6, Hibernate 3.3 para la persistencia de los datos y Seam 2.1.1 como el marco de trabajo integrador. Se espera como resultado que la aplicación de la solución propuesta detecte la presencia de ruido en el registro de eventos, generando modelos que permitan una mejor comprensión y usabilidad para usuarios no expertos en Minería de Procesos.

Palabras claves: *Fuzzy Miner*, Minería de Procesos, *plugin*, registro de eventos, Sistema de Información Hospitalaria.

¹plugin: es un concepto de la lengua inglesa que puede entenderse como “inserción” y que se emplea en el campo de la informática. Añade una funcionalidad adicional o una nueva característica al software.

ÍNDICE

INTRODUCCIÓN	1
CAPÍTULO 1. FUNDAMENTOS TEÓRICO–METODOLÓGICOS DE LA INVESTIGACIÓN.....	6
1.1. Proceso	6
1.2. Gestión basada en procesos	6
1.3. Clasificación de los procesos	7
1.4. Modelos de procesos	8
1.5. Registro de eventos	9
1.6. Minería de Procesos	9
1.6.1. Tipos de Minería de Procesos.	10
1.6.2. Desafíos de la Minería de Procesos	11
1.7. Herramientas con capacidades de Minería de Procesos	11
1.7.1. Herramientas académicas	12
1.7.2. Herramientas comerciales	12
1.7.3. Resultados del estudio de las herramientas con capacidades de Minería de Procesos.	14
1.8. Plugin Fuzzy Miner	15
1.8.1. Formas en las que se manifiesta el ruido en los procesos	16
1.8.2. Ventajas y limitaciones del plugin Fuzzy Miner	17
1.8.3. Investigaciones de la aplicación de Fuzzy Miner en el sector de salud	17
1.9. Ambiente de desarrollo	18
1.9.1. Metodologías de desarrollo de software	19
1.9.2. Lenguaje de programación	20
1.9.3. Herramientas a utilizar	21
1.9.4. Tecnologías a emplear	22
1.10. Conclusiones del capítulo	25
CAPÍTULO 2. FUZZY MINER: ANÁLISIS Y PROPUESTA DE SOLUCIÓN	26
2.1. Propuesta de solución	26
2.2. Patrones de diseño y arquitectura	27
2.2.1. Patrón arquitectónico Modelo-Vista-Controlador.....	27
2.2.2. Patrón de diseño GRASP	28
2.3. Análisis del plugin Fuzzy Miner de ProM	29
2.3.1. Configuración de métricas	29
2.3.2. Configuración de medición	32
2.3.3. Proceso de minería.....	33

2.4. Procedimiento para la detección de ruido con el plugin Fuzzy Miner para el HIS	36
2.4.1. Fase 1: Determinar las configuraciones necesarias para el HIS	36
2.4.2. Fase 2: Integrar el plugin Fuzzy Miner al HIS	38
2.4.3. Fase 3: Mostrar el modelo desde el HIS	41
2.5. Modelo obtenido con Fuzzy Miner desde el HIS y comprensión del mismo	43
2.6. Conclusiones del capítulo	44
CAPÍTULO 3. VALIDACIÓN DE LA SOLUCIÓN PROPUESTA	45
3.1. Validación	45
3.1.1. Cantidad de actividades	45
3.1.2. Identificación de ruido en actividades	47
3.1.3. Identificación de ruido en aristas	51
3.2. Conclusiones parciales	53
CONCLUSIONES	54
RECOMENDACIONES	55
REFERENCIAS BIBLIOGRÁFICAS	56
ANEXOS	61

ÍNDICE DE FIGURAS Y TABLAS

Figura 1: Ejemplo Proceso Lasaña..... 7

Figura 2: Ejemplo Proceso Espaguetis..... 8

Figura 3: Ejemplo de un registro de eventos. 9

Figura 4: Tipos de técnicas de Minería de Procesos.10

Figura 5: Propuesta de la solución.27

Figura 6: Interacción MVC.....28

Figura 7: Interfaz de ProM para visualizar las métricas unarias.....30

Figura 8: Interfaz de ProM para visualizar las métricas binarias.30

Figura 9: Interfaz de ProM para visualizar las métricas de correlación binaria.....31

Figura 10: Interfaz de ProM para visualizar la distancia máxima de evento y la atenuación.33

Figura 11: Interfaz de ProM de un modelo difuso.34

Figura 12: Ejemplo de filtros del plugin Fuzzy Miner de ProM.34

Figura 13: Interfaz de ProM de la Vista de métricas unarias.....35

Figura 14: Interfaz de ProM de la Vista de métricas binarias.36

Figura 15: Diagrama de clases.....39

Figura 16: Interfaz de Jboss de la jerarquía de clases del componente.....40

Figura 17: Interfaz de Jboss de las instancias de la clase Analisis_Proceso.java.....40

Figura 18: Interfaz del HIS de los parámetros de configuración para el modelado.41

Figura 19: Ejemplo de la personalización de Fuzzy Miner en el HIS.....42

Figura 20: Secuencia principal del proceso con *Fuzzy Miner*.43

Figura 21: Interfaz de la herramienta ProM de selección del plugin Add identities to log.46

Figura 22: Interfaz de la herramienta ProM del plugin Add identities to log que muestra la cantidad de actividades.46

Figura 23: Interfaz del proceso “Solicitar productos” con el plugin Fuzzy Miner desde el HIS.....47

Figura 24: Interfaz de la herramienta ProM del proceso “Solicitar productos”.48

Figura 25: Interfaz del proceso “Solicitar productos” con el plugin Fuzzy Miner desde el HIS.....49

Figura 26: Interfaz del HIS para la configuración del valor Frecuencia de actividades.....50

Figura 27: Interfaz de actividades con frecuencia menor a 30 por ciento.51

Figura 28: Interfaz del HIS de aristas con ruido en el proceso “Desincorporar producto”.....52

Figura 29: Diagrama de casos de uso Generar modelo de procesos.62

Figura 30: Interfaz de la herramienta ProM para importar registro de eventos.67

Figura 31: Interfaz de la herramienta ProM del proceso “Solicitar productos” con todas las métricas activas.....68

Figura 32: Interfaz de la herramienta ProM para desactivar métricas unarias.	68
Figura 33: Interfaz de la herramienta ProM del proceso “Solicitar productos” desactivando las métricas unarias.	69
Figura 34: Interfaz de la herramienta ProM para desactivar métricas binarias.	69
Figura 35: Interfaz de la herramienta ProM del proceso “Solicitar productos” desactivando las métricas binarias.	70
Figura 36: Interfaz de la herramienta ProM para desactivar métricas de correlación binaria.	70
Tabla 1: Ejemplos de productos de software con capacidades de Minería de Procesos.	14
Tabla 2: Interpretación de la presencia de ruido con datos obtenidos por el plugin Fuzzy Miner desde el HIS.	49
Tabla 3: Ejemplo de interpretación de ruido en aristas del proceso “Desincorporar Producto” del HIS.	52
Tabla 4: Requisito funcional del componente para la detección de variabilidad en procesos hospitalarios.	62
Tabla 5: Escenario Generar modelo de procesos.	65
Tabla 6: Descripción de las variables.	66
Tabla 7: Valores para medir frecuencia de aristas.	71
Tabla 8: Valores para medir la correlación de aristas.	71

INTRODUCCIÓN

El grado de avance que alcanzan las Tecnologías de la Información y las Comunicaciones (TIC) propicia el desarrollo informático de las organizaciones. La mayoría de las tareas de empresas, organismos e instituciones son ejecutadas y controladas por sistemas de información por ser un conjunto de componentes interrelacionados que recolectan, procesan, almacenan y distribuyen información (Laudon, K.C. y Laudon, J.P., 2012) para apoyar la toma de decisiones y el control de una organización.

Con el fin de obtener resultados eficientes, constituye una prioridad mejorar la administración y desempeño organizacional a través de la incorporación de la gestión o Enfoque basado en procesos. El Enfoque basado en procesos es un principio de gestión básico y fundamental para la obtención de resultados. Este principio sostiene que “un resultado se alcanza más eficientemente cuando las actividades y los recursos se gestionan como un proceso” (ISO 9001, 2008). Entiéndase por proceso la actividad o conjunto de actividades que utiliza recursos, y que se gestiona con el fin de permitir que los elementos de entrada se transformen en resultados.

En este sentido, las instituciones del sector de la salud, se han ido modernizando con los sistemas y las tecnologías de información que han surgido, permitiéndoles avanzar en la prestación de servicios, sin embargo esto va generando considerables volúmenes de datos derivados de las actividades de los procesos clínicos que necesitan ser analizados para que su aporte sea verdaderamente significativo. Como alternativa para el análisis de procesos surge la Minería de Procesos.

La Minería de Procesos posibilita entender cómo son ejecutados en realidad los procesos en el sistema (Van Doremán, 2012), pretende descubrir, monitorear y mejorar (Salas y Rozinat, 2013) procesos reales mediante la extracción de conocimiento de los registros de eventos disponibles en los sistemas información. Los *logs*² son utilizados para realizar los tres tipos de Minería de Procesos existentes: descubrimiento, conformidad y mejora, siendo el descubrimiento de procesos la técnica más investigada.

Sin embargo, la tarea de descubrimiento de procesos puede dificultarse si aparecen aspectos como el ruido (Salas y Rozinat, 2013). Para esta investigación se asume la definición de ruido: “Comportamiento reflejado en las trazas y que rara vez ocurre, que es excepcional o poco frecuente, es decir, que no se corresponde con el comportamiento típico observado en el proceso”

² logs: nace del término en inglés log, que significa registro.

propuestas por Aalst (Van der Aalst, W.M.P., 2011a), por tal motivo detectar ruido constituye un reto y una de las líneas de investigación de la Minería de Procesos.

Entre los autores más destacados en investigaciones acerca de Minería de Procesos en el área de salud, se puede citar a Will van der Aalst, quien destaca los beneficios y desafíos (Van der Aalst, W.M.P. et al., 2011) al aplicar esta disciplina en un sector crítico de la sociedad, como lo es la atención sanitaria. Además, en investigaciones realizadas, se evidencia que es una tecnología novedosa y efectiva para la obtención y análisis de procesos hospitalarios, su uso en casos reales de procesos de ginecología y oncología (Mans et al., 2009) permitió optimizar, a partir de un gráfico de puntos, la trayectoria de los pacientes por el proceso de atención. Tras descubrir horarios de mayor afluencia en el área de Emergencias (Mans et al., 2008), permitió controlar y destinar recursos a este servicio.

Webster (Webster, 2011) realiza una investigación sobre la gestión de los procesos de negocios en las Hojas Clínicas Electrónicas (EHR, por sus siglas en inglés), donde se demuestra la importancia de la Minería de Procesos como un nuevo método para mejorar los procesos de cuidado a pacientes. Por su parte Maruster (Maruster, 2011) propone en su investigación sobre las Redes de Flujo (*Workflow Nets*, por su nombre en inglés), la aplicación de las Redes de Petri para el modelado de información hospitalaria.

Las técnicas de Minería de Procesos son las encargadas de extraer información no trivial y útil de los registros de eventos almacenados por sistemas de información (Yzquierdo, 2012). Una de las técnicas de la Minería de Procesos es la Minería Difusa (*Fuzzy Miner*, por su nombre en inglés) la que es capaz de detectar y manejar con éxito la presencia de ruido.

Con la aplicación de esta técnica se obtiene un modelo basado en gráficos capaz de proporcionar una vista de alto nivel de un proceso, con la abstracción de los detalles no deseados. Permite la obtención de modelos con un adecuado balance pre-ajuste y sobre-ajuste respecto a los registros de eventos y los resultados son obtenidos de forma inmediata, lo cual agiliza la obtención de los resultados y proporciona una gran versatilidad a la hora de realizar análisis y toma de decisiones.

En el Centro de Informática Médica (CESIM) de la Universidad de las Ciencias Informáticas (UCI) se desarrolla un Sistema de Información Hospitalaria (HIS, por sus siglas en inglés). Sobre el sistema se desarrolló un componente para la extracción y transformación de los sucesos y las acciones realizadas durante su ejecución, el mismo permite extraer registros de eventos de los diferentes procesos del sistema. Sin embargo para su análisis es necesario hacer uso de

herramientas externas al dominio hospitalario como por ejemplo ProM (Van der Aalst, W.M.P. et al., 2009) herramienta académica que soporta una variedad de técnicas de Minería de Procesos.

En investigaciones realizadas sobre los procesos del sistema en cuestión y con apoyo de la herramienta antes mencionada, se demostró que existen eventualidades en la ejecución de los mismos, como el comportamiento anormal e infrecuente en el registro de eventos (Ruido). En este sentido, es importante que el sistema cuente con la forma de obtener puntos de vistas cualitativos y cuantitativos sobre la existencia y detección de ruido en sus procesos, así como que implemente un procedimiento para analizar los elementos que representan ruido en los procesos ejecutados para que el usuario sea consciente de las actividades que constituyen ruido y pueda realizar un correcto análisis del modelo.

Con las investigaciones se ha comprobado que los modelos generados por la herramienta ProM no son de la comprensión de los usuarios no expertos en el área de Minería de Procesos y las herramientas que aplican otros sectores son privativas y/o nativas en sistemas automatizados. Al no identificar el ruido pueden obtenerse modelos demasiado complejos o que no representen adecuadamente el proceso, puede obviar en la representación del proceso, comportamientos que son infrecuentes por definición lo que limita su análisis.

La presencia de ruido tiene un impacto estructural considerable sobre el modelo obtenido, debido a que puede influir en el tipo de patrón de control de flujo identificado o sugerido por la técnica. En las secciones del registro de eventos donde está presente, como consecuencia puede significar que ocurren actividades anómalas en el proceso que atenten contra su correcto funcionamiento dado que brindan información acerca de las posibles violaciones del modelo. La presencia de desviaciones en el modelo constituye un indicador de posibles fraudes o violaciones en las políticas lo cual puede ser útil a las organizaciones en la realización de auditorías.

Asumir que faltan determinadas evidencias de la ejecución del proceso, son estimaciones que sólo pueden ser confirmadas a partir del contexto de ejecución particular del proceso analizado y la técnica empleada para su identificación. El usuario debe poder decidir cuáles comportamientos deben ser considerados como ruido y descartarlos del modelo resultante, sin embargo, se ve limitado únicamente a contar con la fuente de información y no con la forma para detectar eventualidades.

Después de haber realizado un análisis de la **situación problemática** y basado en (Sampieri, 1997; Álvarez de Zayas, 1999), se identifica en el marco teórico el siguiente **problema a resolver**:

¿Cómo detectar la presencia de ruido en los procesos del Sistema de Información Hospitalaria del CESIM a partir de su registro de eventos?

Se define como **objeto de estudio** gestión basada en procesos hospitalarios y como **campo de acción** la técnica *Fuzzy Miner* de Minería de Procesos.

El **objetivo general** es desarrollar una personalización del *plugin Fuzzy Miner* de ProM para el Sistema de Información Hospitalaria del CESIM, que permita la detección de ruido en los procesos, a partir de su registro de eventos.

Como **posible resultado** a obtener mediante la realización de este trabajo se propone:

Obtener una personalización del *plugin Fuzzy Miner* de ProM que permita detectar el ruido en los procesos del Sistema de Información Hospitalaria del CESIM, para su análisis y toma de decisiones por parte del personal administrativo.

Para alcanzar el objetivo planteado se proponen las siguientes **tareas de la investigación**:

1. Caracterizar el estado actual de la Minería de Procesos en el mundo para determinar sus aspectos más relevantes.
2. Diagnosticar el estado actual del desarrollo, aplicación y modelación de la Minería de Procesos en el campo de salud para conocer su situación real en el dominio hospitalario.
3. Elaborar el marco teórico metodológico referente a la detección de ruido en procesos de negocio, la Minería de Procesos y los conceptos asociados a la gestión basada en procesos hospitalarios para realizar un análisis e identificar las herramientas que permita personalizar la técnica.
4. Personalizar el *plugin Fuzzy Miner* de ProM sobre el Sistema de Información Hospitalaria del CESIM, para la detección de ruido en sus procesos.
5. Validar los resultados obtenidos con la herramienta ProM 6.4 para comprobar la solución propuesta en cuanto a sus especificaciones.

Como métodos de investigación se utilizó el Histórico-lógico, que permitió constatar teóricamente cómo ha evolucionado, siguiendo la lógica de su desarrollo, la técnica *Fuzzy Miner* de Minería de Procesos desde su surgimiento, el Análisis-síntesis permitió el análisis de las teorías y documentos relacionados con la investigación para la extracción de los elementos más importantes que se relacionan con la técnicas *Fuzzy Miner* y de las herramientas utilizadas en la aplicación de Minería de Procesos.

La Modelación permitió la representación de la propuesta de solución. La Entrevista posibilitó la recolección de información a través de conversaciones planificadas con especialistas y miembros del grupo de investigación de Minería de Procesos en la UCI (**Anexo 1**), y por último, se utilizó la Observación como método para validar la propuesta de solución.

El documento está estructurado en tres capítulos, que incluyen elementos teóricos sobre Minería de Procesos y de la técnica *Fuzzy Miner*, el procedimiento para desarrollar la personalización del plugin, su aplicación y validación en el sistema.

El **Capítulo 1. Fundamentos teórico-metodológicos de la investigación**, aborda los conceptos fundamentales sobre Minería de Procesos: tipos, técnicas, clasificación de los procesos, modelo de procesos, desafíos y otros que se necesitan para comprender cuándo el ruido está presente en los registros de eventos y cómo se trata, dígame, formas en que se manifiesta el ruido en los procesos y especificaciones generales sobre la técnica *Fuzzy Miner*; así como las herramientas a utilizar. Además se analizan varias propuestas de diferentes autores sobre el uso de la técnica *Fuzzy Miner* aplicada al sector, o área de salud.

En el **Capítulo 2. Fuzzy Miner: análisis y propuesta de solución**, se describen los patrones de diseño y arquitectura a utilizar, se analiza el *plugin Fuzzy Miner* de ProM, se presenta la propuesta de solución y se describe el procedimiento a seguir para la personalización del *plugin* sobre el HIS del CESIM que permitirá detectar la presencia de ruido en sus registros de eventos.

Por último, en el **Capítulo 3. Validación de la solución propuesta** se describirán las validaciones realizadas a la personalización propuesta como complemento del Sistema de Información Hospitalaria, haciendo uso de la personalización del *plugin Fuzzy Miner* para el HIS se evalúan un conjunto de procesos y se analizan los resultados obtenidos.

CAPÍTULO 1. FUNDAMENTOS TEÓRICO–METODOLÓGICOS DE LA INVESTIGACIÓN

En el presente capítulo se abordan conceptos fundamentales sobre Minería de Procesos para facilitar la comprensión del alcance de la investigación. Se abordan temas como: gestión basada en procesos, clasificación de los procesos, tipos de Minería de Procesos, desafíos de la Minería de Procesos y otros conceptos necesarios para saber cuándo se está en presencia del ruido y cómo tratarlo, dígame, formas en que se manifiesta y especificaciones generales sobre la técnica *Fuzzy Miner*. Posteriormente se analizan investigaciones en el campo de la salud donde se ha aplicado la técnica, sus beneficios y ventajas y por último se expone una visión general del lenguaje de programación, tecnologías y herramientas a utilizar para cumplir el objetivo de la investigación.

1.1. Proceso

Según la ISO (ISO 9001, 2008) se define un proceso como “una actividad o un conjunto de actividades que utiliza recursos, y que se gestiona con el fin de permitir que los elementos de entrada se transformen en resultados” y el Soporte a Procesos Empresariales (*Enterprise Process Support*, EPS por sus siglas en inglés) lo plantea como “conjunto de tareas de trabajo interrelacionadas que se inician en respuesta a un evento y logra un resultado específico para el cliente del proceso”. Los autores de este trabajo asumen como definición de proceso para esta investigación, la presentada por la norma ISO 9001.

1.2. Gestión basada en procesos

Con el fin de obtener resultados eficientes constituye una prioridad mejorar la administración y desempeño organizacional a través de la incorporación de la gestión o enfoque basado en procesos. El Enfoque Basado en Procesos es un concepto que aparece en los años 90 del pasado siglo como principio de gestión básico y fundamental para la obtención de resultado, el mismo permite dar seguimiento a los procesos, permitiendo detectar errores, redundancias y gestionar los procesos interrelacionados.

La Gestión de Procesos de Negocio (*Business Process Management*, BPM por sus siglas en inglés) es la disciplina que combina conocimiento sobre tecnología de información y conocimiento sobre las ciencias de gestión y lo aplica en conjunto a los procesos de negocio operacionales (Van der Aalst, W.M.P. et al., 2011), permite diseñar, ejecutar, controlar y analizar procesos operacionales cuyo objetivo es mejorar el desempeño y la optimización de los procesos de negocio de una organización. A partir de BPM surgen los Sistemas de Administración de Procesos

de Negocios (*Business Process Management Systems*, BPMS por sus siglas en inglés), para permitir el entendimiento de la organización y su desempeño a todos los interesados.

1.3. Clasificación de los procesos

Los procesos, según (Van der Aalst, W.M.P., 2011a), se clasifican en dos categorías: estructurados (Proceso Lasaña) y desestructurados (Proceso Espaguetis). Cuando en los registros de eventos se muestra que los procesos se caracterizan por una manera de ejecutarse controlada y que poseen una clara estructura, dichos procesos son estructurados, para esta clasificación es fácil la aplicación de técnicas de Minería de Procesos pues los procesos están organizados, además las mejoras son pequeñas.

Por otra parte, cuando los registros de eventos tienen actividades ejecutadas con diversos órdenes, variando el momento en el que se ejecutaron, se cambian los autores, se evidencian muchas relaciones o variaciones entre las actividades, se está en presencia de los procesos desestructurados. En la **Figura 1** se muestra un ejemplo de un Proceso Lasaña en el que se evidencia lo estructurado del proceso.

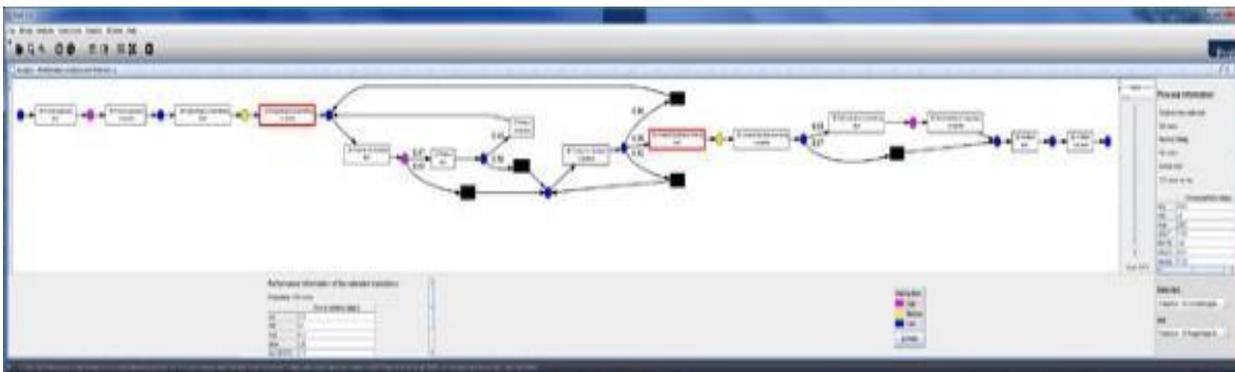


Figura 1: Ejemplo Proceso Lasaña.

Fuente: (Van der Aalst, W.M.P. 2011a).

En la **Figura 2** se muestra un ejemplo de un Procesos Espaguetis en el que se ponen de manifiesto muchas actividades relacionadas entre sí, lo cual dificulta el análisis del modelo.

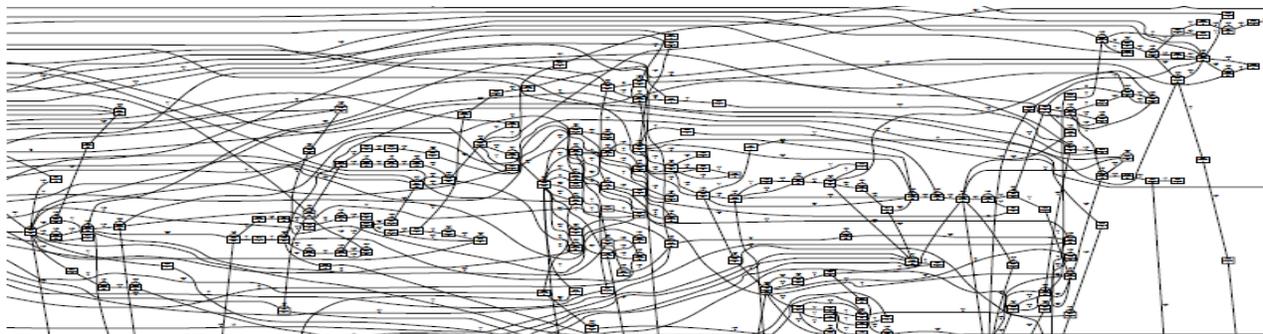


Figura 2: Ejemplo Proceso Espaguetis.

Fuente: (Van der Aalst, W.M.P. y Günther, Christian W., 2007)

Se concluye que la clasificación de los procesos dependerá en todo momento del orden en que se ejecuten y de las cantidades de actividades relacionadas entre sí.

1.4. Modelos de procesos

Muchas veces los procesos o subprocessos son difíciles de comprender, son complejos, confusos, muy grandes e incluso pueden contener múltiples puntos de contacto entre sí que dificultan su análisis y usabilidad. Por otra parte un modelo puede dar la oportunidad de organizar un proceso de manera que sea muy fácil su comprensión pues es una representación de una realidad. Modelar es desarrollar una descripción lo suficientemente buena de actividades, tareas o procesos en general.

Cuando un proceso es modelado en una representación gráfica se aprecia con facilidad las interrelaciones existentes entre distintas actividades se pueden analizar las actividades independientemente, es fácil distinguir entre aquellas que aportan valor añadido a las que no lo hacen, se logra definir los puntos de contacto con otros procesos así como identificar los subprocessos comprendidos. Al mismo tiempo los problemas existentes pueden ponerse de manifiesto claramente dando la oportunidad al inicio de acciones de mejora.

Por tanto, con la información anterior los autores de la presente investigación definen que un modelo de proceso es: la representación gráfica a partir de técnicas y herramientas especializadas de una realidad compleja, con el fin de facilitar su comprensión, comportamiento y análisis.

1.5. Registro de eventos

Un registro de eventos (**Figura 3**) es la colección de eventos³ utilizados como entrada para la Minería de Procesos (Van der Aalst, W.M.P. et al., 2011). La mayoría de los sistemas de información actuales poseen un registro de trazas, las trazas son las ejecuciones de los procesos.

case id	event id	properties				
		timestamp	activity	resource	cost	...
1	35654423	30-12-2010:11.02	register request	Pete	50	...
	35654424	31-12-2010:10.06	examine thoroughly	Sue	400	...
	35654425	05-01-2011:15.12	check ticket	Mike	100	...
	35654426	06-01-2011:11.18	decide	Sara	200	...
	35654427	07-01-2011:14.24	reject request	Pete	200	...
2	35654483	30-12-2010:11.32	register request	Mike	50	...
	35654485	30-12-2010:12.12	check ticket	Mike	100	...
	35654487	30-12-2010:14.16	examine casually	Pete	400	...
	35654488	05-01-2011:11.22	decide	Sara	200	...
	35654489	08-01-2011:12.05	pay compensation	Ellen	200	...

Figura 3: Ejemplo de un registro de eventos.

Fuente: (Van der Aalst, W.M.P. 2011a).

1.6. Minería de Procesos

El punto de partida de esta tecnología es BPM, que es la base para los Sistemas de Administración de Procesos de Negocios que generan la información sobre la que trabaja la Minería de Procesos. La Minería de Procesos es una disciplina que tiene por objeto la detección y extracción automática de la información valiosa de registros de eventos (Van der Aalst, W.M.P. et al., 2007; Dustdar, Hoffmann y Van der Aalst, W.M.P., 2005; Rozinat y Van der Aalst, W.M.P. 2008b) para captar el proceso de negocio a medida que se ejecuta y su definición es la siguiente:

La Minería de Procesos es una disciplina de investigación relativamente joven que se ubica entre la inteligencia computacional y la Minería de Datos⁴, por una parte, y la modelación y análisis de procesos, por otra. La idea de la Minería de Procesos es descubrir, monitorear y mejorar los procesos reales a través de la extracción de conocimiento de los registros de eventos disponibles en los actuales sistemas de información. Su aplicación ayuda a identificar cuellos de botella, anticipar problemas, registrar violaciones de políticas, recomendar contramedidas y simplificar procesos (Van der Aalst, W.M.P., 2011a).

³eventos: son acciones almacenadas en el registro, por ejemplo, el inicio, conclusión o cancelación de una actividad para una instancia particular de un proceso.

⁴Minería de Datos: análisis de conjuntos de datos (a menudo grandes) para encontrar relaciones inesperadas y para resumir los datos de manera que proporcionen nuevos entendimientos (Van der Aalst, W.M.P. et al. 2011).

1.6.1. Tipos de Minería de Procesos.

Las técnicas de Minería de Procesos permiten extraer información no trivial y útil de los registros de trazas almacenados por sistemas de información (Yzquierdo, 2012). Existen tres tipos fundamentales de técnicas de Minería de Procesos, el descubrimiento de procesos, la verificación de conformidad y el mejoramiento de modelos, como se muestra en la **Figura 4**.



Figura 4: Tipos de técnicas de Minería de Procesos.

Fuente: adaptado de (Van der Aalst, W.M.P., 2011a).

En el descubrimiento de procesos se usa un registro de eventos para producir un modelo. El descubrimiento de procesos es la técnica de Minería de Procesos más atendida dentro de esta área y se han diseñado diversos algoritmos con diferentes enfoques para ella. Entiéndase por algoritmo de descubrimiento una función que mapea un registro de eventos hacia un modelo de proceso (Van der Aalst, W.M.P., 2011a). Para muchas organizaciones es sorprendente ver que las técnicas existentes son realmente capaces de descubrir los procesos reales meramente basadas en las muestras de ejecución en los registros de eventos. Entre los variados algoritmos existentes se encuentran:

- Fuzzy Mining (Van der Aalst, W.M.P. y Günther 2007).
- Alpha Mining (Van der Aalst, W.M.P., 2011a).
- Heuristic Miner (Van der Aalst, W.M.P., 2011a; Weijters y Ribeiro, 2011; Van der Aalst, W.M.P., Medeiros y Weijters, 2007).
- Genetic Mining (Bratosin, Sidorova y Van der Aalst, W.M.P., 2010; Van der Aalst, W.M.P., Medeiros y Weijters, 2007; Medeiros, A.K., Weijters y Van der Aalst, W.M.P., 2005).

Para la Verificación de Conformidad se necesitan un registro de eventos y un modelo como entrada. La salida consiste en información de diagnóstico mostrando las diferencias y elementos en común entre el modelo y el registro de eventos. Aquí se compara un proceso existente con un

registro de eventos del mismo proceso para verificar si la realidad, según el registro, es equivalente al modelo y viceversa.

La Mejora de los Modelos busca extender o mejorar un modelo de proceso existente con la información del proceso real almacenado en un registro de eventos. También necesitan un registro de eventos y un modelo como entrada. La salida es un modelo mejorado o extendido.

1.6.2. Desafíos de la Minería de Procesos

Los desafíos de los que se quiere ocupar la Minería de Procesos se plantean y analizan en (Van der Aalst, W.M.P. y Weijters, 2004; Van der Aalst, W.M.P. et al., 2011). En (Van der Aalst, W.M.P. y Weijters, 2004) se recogen varios de los problemas desafiantes que presenta la Minería de Procesos como son las tareas ocultas, tareas duplicadas, los bucles, el tiempo, el ruido, entre otros problemas en la aplicación práctica de Minería de Procesos.

En (Van der Aalst, W.M.P. et al., 2011) estos desafíos, se resumen y actualizan pues se necesita superar varios obstáculos tales como la distribución en varias fuentes, que a menudo están “centrados en objetos” más que “centrados en procesos”, otras veces incompletos. Además el registro de eventos puede contener datos atípicos o eventos en diferentes niveles de granularidad. Esta investigación se centra en el desafío de cómo tratar con el ruido en los registros de eventos, sin embargo, el desafío de mejorar la usabilidad y entendimiento de los modelos por parte de los usuarios no expertos es un tema a considerar en la implementación de la propuesta de solución.

1.7. Herramientas con capacidades de Minería de Procesos

Luego de realizar una búsqueda de herramientas asociadas a la investigación tanto en el ámbito internacional como nacional para el análisis de modelos de procesos enfocados en la forma que muestran y manejan el ruido, destacaron en el ámbito internacional las herramientas ProM y Disco como más difundidos, ProM desarrollada por la Universidad Tecnológica de Eindhoven y Disco desarrollada por la compañía Fluxicon.

Existen además otros sistemas desarrollados con fines comerciales como son: ARIS Process Performance Manager, Discovery Analyst, QPR ProcessAnalyzer, entre otros y sus mayores aplicaciones radican en las auditorías a sistemas y el control del flujo de información. En el ámbito nacional solo existe documentación publicada de la Minería de Variantes, una técnica de Minería de Procesos desarrollada para el marco de trabajo ProM que está basada en algoritmos para el diagnóstico de variantes de procesos de negocio a partir de registros de eventos con ruido y/o ausencia de información.

1.7.1. Herramientas académicas

ProM

La herramienta ProM de código abierto y distribuida gratuitamente ha sido el estándar impuesto para la Minería de Procesos durante la última década. Permite el proceso de descubrimiento, la comprobación de la conformidad, análisis de redes sociales, la minería de organización, la minería de decisión (Van der Aalst, W.M.P., 2011a; Verbeek, 2010). ProM es el *software* que concentra la mayor cantidad de herramientas y algoritmos que existen en Minería de Proceso.

Minería de Variantes

Herramienta informática basada en algoritmos para el diagnóstico de variantes de procesos de negocio. El objetivo perseguido por la Minería de Variantes es construir variantes de proceso a partir de diferentes descomposiciones en subprocesos considerando patrones de control de flujo. Las variantes de descomposición obtenidas se muestran en un Árbol de Variantes, asociado a un Perfil de Diagnóstico, además esta técnica posibilita la identificación de características del registro de eventos como el ruido y la ausencia de información. (Pérez, 2014).

Disco

Disco es una aplicación completa de Minería de Procesos desarrollada por Fluxicon en 2009, es la herramienta más usada dentro de las privativas, posee una licencia gratuita limitada con fines académicos. Implementada con el objetivo de ser una herramienta profesional para el apoyo a las organizaciones en el control de sus procesos. Es totalmente compatible con las herramientas académicas ProM 5 y 6. La minería de Disco se basa en *Fuzzy Miner* pero se ha desarrollado en muchas maneras. *Fuzzy Miner* fue el primer algoritmo de minería en introducir el "mapa difuso" a la Minería de Procesos. Para Disco, se han utilizado las ventajas de *Fuzzy Miner* combinado con la experiencia de la propia práctica y pruebas de usuario. A pesar de que se basa en el marco de *Fuzzy Miner* se ha desarrollado un nuevo conjunto de métricas y estrategias de modelado de procesos haciendo efectiva la minería de Disco como una próxima generación de *Fuzzy Miner*.

1.7.2. Herramientas comerciales

ARIS Process Performance Manager

ARIS Process Performance Manager (ARIS PPM) es una herramienta desarrollada por Software AG la cual permite a las empresas supervisar y analizar el rendimiento y la estructura de sus procesos empresariales. El ARIS PPM cuenta con funcionalidades como: supervisión de los valores de destino y del estado actual; medición del rendimiento de los procesos; análisis de

comunicación y actividades de departamentos y otras unidades organizativas (Software AG, 2011), entre otras funcionalidades que posibilitan a los usuarios visualizar automáticamente modelos de procesos reales a partir de los datos del sistema y de este modo realizar comparaciones de los objetivos frente a los valores reales de los flujos de trabajo.

Discovery Analyst

Discovery Analyst es una herramienta desarrollada por StereoLOGIC, extrae los procesos de negocio de las aplicaciones en tiempo real con el objetivo de crear modelos de procesos. Cuenta con funcionalidades como: la Grabación cuyo objetivo es grabar lo que un usuario ha presionado para llegar a una pantalla en particular. También cuenta con la Visualización donde los usuarios pueden simular las actividades de procesos de negocio. Con la Generación de procesos de negocio se obtienen automáticamente mapas de procesos de negocio en diferentes formatos para su posterior mejora y transformación.

Además cuenta con la funcionalidad de Medición que captura el tiempo real dedicado por los empleados que realizan el proceso de negocios. La Comparación permite diferenciar entre los procesos descubiertos y por último la Generación de documentación que no es más que la documentación estructurada y trazable, incluye mapas de procesos, etapas de proceso y pantallas de la aplicación asociada, se puede generar de forma automática.

QPR ProcessAnalyzer

QPR ProcessAnalyzer es un *Software Comercial Automatizado de Descubrimiento de Procesos* el cual permite a su organización acelerar las iniciativas de gestión de procesos de negocio, reduciendo el tiempo y los costos involucrados con las acciones de mejora de procesos. A partir del uso de los datos almacenados en los sistemas de negocio operativos, *QPR ProcessAnalyzer* muestra exactamente cómo sus procesos se ejecutan en realidad y permite analizar desde múltiples ángulos, así como profundizar en los casos individuales (QPR Software Oyj, 2011).

Entre las funcionalidades con que cuenta están: realizar análisis basados en hechos a partir de fracciones de tiempo, descubrir modelos de procesos para el análisis de rendimiento que permite analizar cuáles son las causas de las variaciones de procesos, descubre y elimina los cuellos de botella del proceso, mejora la eficiencia operativa y reduce los costos a través del rendimiento de los procesos y por último optimiza la toma de decisiones de forma rápida con el descubrimiento de procesos automatizados.

1.7.3. Resultados del estudio de las herramientas con capacidades de Minería de Procesos

Existen cuantiosos productos de *software* distribuidos por diferentes empresas bajo licencias propietarias que cuentan con capacidades de Minería de Procesos sin embargo el estudio realizado de los sistemas anteriores y características fundamentales (**Tabla 1**) ha permitido apreciar que la mayoría de los sistemas solo son factibles en un estrecho dominio que no incluye al hospitalario, además se han desarrollado bajo tecnologías privativas y a la medida.

En el caso de ProM es libre y además una herramienta genérica pero aun así externas al dominio hospitalario, la Minería de Variantes igualmente libre pero desarrollada para el marco de trabajo ProM. Además todos los sistemas estudiados son aplicaciones de Escritorio y dificultan la posible integración con el HIS por ser una aplicación web.

Adicionalmente ProM no está enfocado a la usabilidad por lo que es recomendable para personas que tienen conocimientos de las técnicas y algoritmos de Minería de Procesos ya que se requiere cierta experiencia para seleccionar las herramientas a usar y para entender e interpretar correctamente los resultados obtenidos, sin embargo es la única herramienta que permite obtener el código por poseer licencia L-GPL. Disco por su parte tiene una licencia con fines académicos y contrario a ProM contiene interfaces amigables e intuitivas pero presenta licencia propietaria, es por ello que los autores de la presente investigación personalizarán el *plugin Fuzzy Miner* de ProM.

Tabla 1: Ejemplos de productos de software con capacidades de Minería de Procesos.

Herramientas/ Características	Licencia	Genérico	Comercial	Tipo de software	Descripción
ProM	Libre, distribuido en partes. Paquete ProM core con licencia GPL. Paquetes ProM plugins bajo licencia L-GPL.	Sí	No	Escritorio	Framework extensible que soporta una variedad de técnicas de Minería de Procesos en forma de plugin.
Disco	Propietario pero posee una licencia gratuita limitada con fines académicos.	Sí	Sí	Escritorio	Aplicación completa de Minería de Procesos.

ARIS Process Performance Manager	Propietario	No	Sí	Escritorio	Plataforma de Minería de Procesos e Inteligencia Empresarial.
Discovery Analyst	Propietario	No	Sí	Escritorio	Extrae los procesos de negocio de las aplicaciones en tiempo real.
QPR ProcessAnalyzer	Propietario	No	Sí	Escritorio	Herramienta para automatizar el descubrimiento de procesos de negocio y QPR ProcessAnalysis, un servicio basado en las citas.
Minería de Variantes	Libre, licencia L-GPL.	Sí	No	Escritorio	Técnica para el diagnóstico de variantes de proceso de negocio para el marco de trabajo ProM.

1.8. Plugin Fuzzy Miner

El *plugin Fuzzy Miner* está basado en el algoritmo *Fuzzy Miner* creado por Günther cofundador de Fluxicon en 2007 y está basado en medidas de significación⁵ y correlación⁶ para visualizar el comportamiento en los registros de eventos. Es el primer algoritmo que se ocupa directamente de los problemas de números grandes de actividades y comportamiento altamente no estructurado.

Según (Günther, 2009) está caracterizada por la presencia de dos tipos de nodos; nodos primitivos que se refieren a una tarea y nodos que se refieren a un conjunto de tareas o clúster (*cluster*, por su nombre en inglés), considerando que cada tarea pertenece a un único nodo y es empleada en el diagnóstico del registro de eventos donde posibilita realizar análisis preliminares al descubrimiento.

⁵significación: importancia de cada una de las actividades con respecto a los demás.

⁶ correlación: es medida por el tiempo de ejecución de dos actividades.

Utiliza similitudes con la forma de representar la información en la cartografía aplicando conceptos como: abstracción, personalización, agregación y énfasis. La agregación se refiere a limitar la cantidad de elementos, nodos y relaciones que se muestran. En su lugar varios elementos forman *cluster* encapsulando la información y facilitando su visualización. La cantidad de información que se muestra es controlada mediante el empleo de umbrales de permisibilidad.

La abstracción consiste en omitir de la visualización la información que en determinado contexto es insignificante. En determinados componentes se destacan características como color, contraste, saturación y tamaño para facilitar la visualización de la información relevante. Por su parte la personalización permite configurar la información que se muestra de acuerdo a determinados criterios (Van der Aalst, W.M.P. y Günther, 2007).

Es una técnica de descubrimiento y este algoritmo es bastante aceptado como técnica de exploración de datos. Es aconsejable utilizar esta técnica cuando tenga los datos de registros complejos y no estructurados o cuando se quiere simplificar el modelo de una manera interactiva. Con su aplicación se obtiene un modelo basado en gráficos capaz de proporcionar una vista de alto nivel de un proceso, con la abstracción de los detalles no deseados.

Fuzzy Miner (Van der Aalst, W.M.P. y Günther, 2007) haciendo uso de umbrales de frecuencia configurables por el usuario descarta del modelo el ruido, sin embargo no considera la ausencia de información en el registro de eventos lo cual puede conducir a representaciones incorrectas en el modelo que ofrece. Tampoco no identifica los patrones de control de flujo presentes. Al no identificar los patrones se dificulta la ubicación en contexto de las anomalías y patrones comunes que ofrecen, limitando la comprensión en el ámbito pragmático.

1.8.1. Formas en las que se manifiesta el ruido en los procesos

Es posible detectar tres formas en las que se manifiesta el ruido en los procesos:

1. Eventos del sistema que no tienen nada que ver con el proceso y que se incluyen en el registro de eventos. La gran cantidad y los cortos intervalos de tiempo en el que estos eventos suceden crean líneas verdes en lugar de un número apreciable de puntos.
 2. Una sola actividad que tuvo lugar mucho antes que el resto de las actividades en el registro de eventos.
 3. Una brecha en el registro donde en un determinado período no hay eventos registrados.
- Dado que estas tres formas de ruido son en realidad ocurrencias distintas en los registros de eventos, son definidos los siguientes términos para tratarlos:

- a) Los casos incompletos son rastros en un registro de eventos (ausencia de eventos de inicio y fin).
- b) El comportamiento excepcional es comportamiento de baja frecuencia que difiere de la mayoría de las trazas.
- c) El ruido del sistema como acontecimientos que se registran incorrectamente debido a un error del sistema o que no tienen nada que ver con el proceso realizado.

1.8.2. Ventajas y limitaciones del plugin Fuzzy Miner

Las principales ventajas de *Fuzzy Miner* radican en la capacidad de actuar en ambientes menos estructurados, se suele decir que es capaz de resolver casos “imposibles” por poder modelar procesos muy grandes y complejos. Además, destaca eventos y relaciones relevantes del proceso, cuenta con la “Separación por tipo de nodo”, los nodos primitivos y nodos *cluster* que se identifican de forma distinta para una mejor visualización. La “Significación de traza” muestra las relaciones con mayor significado y se representan con trazas de mayor grosor. Mientras la “Correlación de traza” también se muestra de forma diferente, las relaciones con mayores índices de correlación se representan con colores más oscuros.

Entre sus principales limitaciones se encuentran las iteraciones, las múltiples instancias y el comportamiento dentro de un *cluster* por lo que se requiere de un estudio más detallado. La comparación con otros algoritmos también está limitada en el sentido de que un modelo difuso no puede transformarse a una Red de Petri (De Weerd, et al., 2012) y por tanto no permite una evaluación comparativa con otras técnicas de descubrimiento de procesos. El modelo difuso que genera la técnica no se puede convertir a otros tipos de lenguajes de modelado de proceso pero sí se puede animar para conseguir una sensación dinámica del proceso.

1.8.3. Investigaciones de la aplicación de Fuzzy Miner en el sector de salud

Se han publicado varios artículos de casos de estudio e investigaciones académicas donde se utiliza dicha técnica en el sector de la salud con el fin de ayudar a la toma de decisiones y optimizar sus principales procesos tales como:

- Minería de Procesos en salud. Caso de Estudio modelado de los procesos del área de Emergencias (Orellana y Sánchez, 2014). Se pone de manifiesto la técnica *Fuzzy Miner* para visualizar el comportamiento de la ejecución de los procesos en el Área de Emergencias a partir de la información registrada en un Sistema de Información Hospitalaria.
- *Process Mining in Healthcare Systems: An Evaluation and Refinement of a Methodology* (Van Bart, 2012) donde evidencia la importancia que presenta el uso de las técnicas de Minería de

Procesos para el análisis y presentación de la información relevante de los procesos en el sector de la salud.

- *Process mining in Intensive Care Unit Data* (Lybeshari, 2012) utiliza la técnica *Fuzzy Miner* para comparar los resultados obtenidos en la Minería de Procesos con otras técnicas y chequear si el equipo de la unidad de cuidados intensivos siguen correctamente los procedimientos definidos previamente.
- Modelado y análisis de un proceso quirúrgico mediante técnicas de Minería de Procesos (Rubio et al., 2010) aborda la importancia que presenta usar técnicas de Minería de Procesos para un correcto modelado e implementación de los procesos quirúrgicos.
- *Application of Process Mining in Healthcare. A Case Study in a Dutch Hospital* (Mans et al., 2009) se aplica la técnica *Fuzzy Miner* para la obtención de conocimientos en casos reales de procesos de ginecología y oncología en un hospital holandés demostrando la necesidad e importancia de su aplicación.

A partir de esta investigación se hace evidente que aunque la aplicación de las técnicas de la Minería de Procesos es reciente (Reijers, Van der Aalst, W.M.P y Weijters, 2007; Rozinat, A. et al., 2009), los investigadores de esta área del conocimiento demuestran su utilidad e importancia en el sector de la salud a nivel mundial que tiene una influencia alta en la calidad de vida de los ciudadanos. Entre los análisis de procesos clínicos administrativos más comunes realizados por expertos en minería aplican *Fuzzy Miner* debido a que brinda las posibilidades de descubrir el comportamiento real del sistema y brinda un mapa de procesos donde se puede analizar el flujo básico y las secuencias principales del mismo.

Dentro de los artículos de casos de estudio e investigaciones académicas en el sector de la salud específicamente en áreas como Cuidados Intensivos, Emergencias, procesos quirúrgicos, procesos administrativos, entre otros, que utilizan *Fuzzy Miner* se demuestra la necesidad de aplicar esta tecnología en registros reales, realizar un análisis de las oportunidades que brinda la Minería de Procesos para monitorear, chequear y mejorar procesos hospitalarios lo cual favorecería a la toma de decisiones clínico administrativas. El empleo de las técnicas en el sector de la salud permitirá hacer análisis de tendencias, comportamientos y eventos demográficos con el objetivo de hacer actualizaciones en programas sociales de salud.

1.9. Ambiente de desarrollo

En un ambiente o entorno de desarrollo se muestra la combinación de herramientas que soporta una gran parte de las tareas del desarrollo de un *software*. Las herramientas deben estar bien

integradas para que puedan inter-operar unas con otras. En resumen el ambiente de desarrollo está formado por el conjunto de instrumentos (*hardware*, *software*, procedimientos, y otros) que facilitan o automatizan las actividades de desarrollo. A continuación se describen las metodologías, lenguaje, las herramientas y las tecnologías para materializar la investigación que están establecidas por el Departamento Sistemas de Gestión Hospitalaria para el desarrollo de los proyectos y que además cumplen con las políticas de independencia tecnológica definidas en Cuba.

1.9.1. Metodologías de desarrollo de software

No existe una metodología de software universal, por tanto toda metodología debe ser adaptada a las características de cada proyecto exigiéndose así que el proceso sea configurable. Se decide hacer una variación de la metodología Proceso Unificado Ágil (AUP, por sus siglas en inglés), de forma tal que se adapte al ciclo de vida definido para la actividad productiva de la UCI. Una metodología de desarrollo de software tiene entre sus objetivos aumentar la calidad del software que se produce de ahí la importancia de aplicar buenas prácticas, para ello la UCI se apoya en el Modelo CMMI-DEV v1.3. El cual constituye una guía para aplicar las mejores prácticas en una entidad desarrolladora.

Variación de AUP para la UCI

De las cuatro fases que propone AUP (Inicio, Elaboración, Construcción, Transición) se decide para el ciclo de vida de los proyectos de la UCI mantener la fase de Inicio pero modificando el objetivo de la misma, se unifican las restantes tres fases de AUP en una sola que es Ejecución y se agrega la fase de Cierre (Rodríguez, 2014).

Inicio: Durante el inicio del proyecto se llevan a cabo las actividades relacionadas con la planeación del proyecto. En esta fase se realiza un estudio inicial de la organización cliente que permite obtener información fundamental acerca del alcance del proyecto, realizar estimaciones de tiempo, esfuerzo y costo y decidir si se ejecuta o no el proyecto (Rodríguez, 2014).

Ejecución: En esta fase se ejecutan las actividades requeridas para desarrollar el software incluyendo el ajuste de los planes del proyecto considerando los requisitos y la arquitectura. Durante el desarrollo se modela el negocio, obtienen los requisitos, se elaboran la arquitectura y el diseño, se implementa y se libera el producto (Rodríguez, 2014).

Cierre: En esta fase se analizan tanto los resultados del proyecto como su ejecución y se realizan las actividades formales de cierre del proyecto (Rodríguez, 2014).

Marco de trabajo para la mejora de procesos CMMI

El modelo CMMI⁷ intenta ser un marco de trabajo para la mejora del proceso que sea aplicable en un amplio abanico de compañías. Su versión en etapas permite un desarrollo del sistema de la organización, la gestión de los procesos a valorar y su asignación a un nivel de madurez entre uno y cinco (Chrissis, Konrad y Shrum, 2009). El centro de desarrollo CESIM se encuentra en el nivel dos de madurez el cual tiene asociado las siguientes áreas de procesos:

1. Administración de requisitos (REQM): Gestiona los requerimientos del proyecto y de sus componentes e identifica inconsistencias entre estos requerimientos y el plan y los productos del proyecto.
2. Planificación de proyecto (PP): Establece y mantiene los planes, los cuales definen las actividades del proyecto.
3. Monitoreo y control del proyecto (PMC): Provee la comprensión del progreso del proyecto y la aplicación de medidas correctivas cuando el rendimiento del proyecto se desvía significativamente del plan.
4. Medición y análisis (MA): Desarrolla y mantiene una capacidad de medición que se utiliza en el soporte de gestión de la información.
5. Aseguramiento de la calidad de procesos y productos (PPQA): Provee personal y gestión con comprensión objetiva del proceso y los productos de trabajo asociados.
6. Administración de la configuración (CM): Establece y mantiene la integridad de los productos de trabajo usando identificación, control y estado de configuraciones, así como auditorías.

1.9.2. Lenguaje de programación

Java 1.6

Java es un lenguaje de programación orientado a objetos que fue desarrollado por James Gosling de Sun Microsystems (la cual fue adquirida por la compañía Oracle) y publicado en 1995 como un componente fundamental de la plataforma Java de Sun Microsystems. Es un lenguaje robusto, pues no permite el manejo directo del *hardware* ni de la memoria. La principal característica de Java es la de ser un lenguaje compilado e interpretado. Todo programa en Java ha de compilarse y el código que se genera es interpretado por una máquina virtual. Dentro de sus principales ventajas se encuentra la de ser multiplataforma por tanto, la personalización a desarrollar podrá ser utilizada desde cualquier entorno (propietario, libre).

⁷ CMMI: Abreviatura de la forma inglesa Capability Maturity Model Integration que significa Integración de modelos de madurez de capacidades.

1.9.3. Herramientas a utilizar

Para la realización de este trabajo se utilizan las herramientas descritas a continuación:

ProM 6.4

ProM es un marco extensible que es compatible con una amplia variedad de técnicas de Minería de Procesos en forma de *plugins*, de *software* libre y multiplataforma. Es independiente de la plataforma por ser implementado en Java y puede ser descargado sin ningún costo. Está publicado bajo una licencia de código abierto. (Van der Aalst, W.M.P., 2011a; Verbeek, 2010). Es utilizado en la investigación para validar la veracidad de la solución propuesta.

Entorno Integrado de Desarrollo Eclipse 3.4.2

Eclipse Ganymede 3.4.2 es un entorno de desarrollo integrado de código abierto, portable y multiplataforma. Actualmente es desarrollado por la Fundación Eclipse que fomenta una comunidad de código abierto. Se basa en el uso de módulos (*plugins*), lo cual hace posible el trabajo en múltiples lenguajes de programación como son Java, C++, PHP, Perl y otros. Es el entorno utilizado para desarrollar la personalización por ser el mejor que se adapta a las necesidades existentes por poseer ventajas como: Debugear el código (permite ver la secuencia de ejecución) para detectar errores, además tendrá soporte para lenguajes o script (JavaScript, css, SQL).

Servidor de base de datos PostgreSQL 8.4

PostgreSQL es un sistema de gestión de bases de datos relacional orientado a objetos y libre, su desarrollo es dirigido por una comunidad de desarrolladores que trabajan de forma libre. En *PostgreSQL* se pueden escribir procedimientos almacenados en más de una docena de lenguajes como: Java, Perl, Python, C++ y otros. El tamaño máximo de la base de datos es ilimitado. Entre sus características más distintivas se encuentra que soporta distintos tipos de datos, incluye herencia entre tablas, es Unicode y está disponible para Linux, UNIX en todas sus variantes y Windows 32/64bit.

PgAdmin III 1.14

PgAdmin III es una herramienta de código abierto para la administración de bases de datos PostgreSQL y derivados. Facilita la gestión y administración de bases de datos ya sea mediante instrucciones del lenguaje de consulta estructurado (SQL, por sus siglas en inglés) o con ayuda de

un entorno gráfico. Permite acceder a todas las funcionalidades de la base de datos; consulta, manipulación y gestión de datos, incluso opciones avanzadas como manipulación del motor de replicación Slony-I. Está disponible en más de una docena de lenguajes y para varios sistemas operativos incluyendo Microsoft Windows, Linux, FreeBSD, Mac OSX y Solaris («pgAdmin», 2015). PgAdmin III soporta versiones de servidores 7.3 y superiores.

Visual Paradigm UML 8.0

Visual Paradigm es una herramienta CASE (Ingeniería de *Software* Asistida por Computación). La misma propicia un conjunto de ayudas para el desarrollo de programas informáticos desde la planificación, pasando por el análisis y el diseño, hasta la generación del código fuente de los programas y la documentación (Pressman, 2002). Es multiplataforma y ha sido concebida para soportar el ciclo de vida completo del proceso de desarrollo del *software* a través de la representación de todo tipo de diagramas. Fue diseñado para una amplia gama de usuarios interesados en la construcción de sistemas de *software* de forma fiable a través de la utilización de un enfoque orientado a objetos. Es utilizada para generar el diagrama de clases de la personalización.

1.9.4. Tecnologías a emplear

Como complementos para lograr la materialización de la investigación se utilizan las siguientes tecnologías:

Lenguaje Unificado de Modelado (UML) 2.1

Es el lenguaje de modelado de sistemas de *software* más conocido y utilizado en la actualidad; está respaldado por el OMG (Object Management Group). Es un lenguaje gráfico para visualizar, especificar, construir y documentar un sistema. UML ofrece un estándar para describir un "plano" del sistema (modelo), incluyendo aspectos conceptuales tales como procesos de negocio, funciones del sistema y aspectos concretos como expresiones de lenguajes de programación, esquemas de bases de datos y compuestos reciclados (Wesley y Stevens, 2002). Es el lenguaje de modelado utilizado en la investigación a través del Visual Paradigm para generar el diagrama de clases.

Java Server Faces (JSF) 1.2

JSF es un *framework* de desarrollo basado en el patrón MVC (Modelo Vista Controlador). Desarrollado a través del *Java Community Process*, la tecnología *JavaServer Faces* establece el estándar para la construcción de interfaces de usuario del lado del servidor (Liu, 2011).

Java Platform Enterprise Edition (JavaEE) 5.0

Java versión 5 o JEE 5 es una plataforma de programación para desarrollar y ejecutar *software* de aplicaciones en lenguaje de programación Java con arquitectura de N niveles distribuida (Franky, 2010). Se centra en hacer el desarrollo más fácil y se basa ampliamente en componentes de *software* modulares y se ejecuta sobre un servidor de aplicaciones.

Java Persistence API (JPA)

Java Persistence API, más conocida por sus siglas JPA, es la API de persistencia desarrollada para la plataforma Java EE e incluida en el estándar EJB3 (*Enterprise JavaBeans*). El objetivo que persigue el diseño de esta API es no perder las ventajas de la orientación a objetos al interactuar con una base de datos, como sí pasaba con EJB2, y permitir usar objetos regulares (conocidos como POJOs).

Hibernate 3.3

Hibernate es una herramienta de mapeo objeto relacional. Utiliza para esto archivos declarativos (XML). Es una tecnología de *software* libre. Como la mayoría de las herramientas de su tipo, busca solucionar el problema de la diferencia entre los dos modelos de datos coexistentes en una aplicación: el usado en la memoria de la computadora (orientación a objetos) y el usado en las bases de datos (modelo relacional), *Hibernate* convertirá los datos entre los tipos utilizados por Java y los definidos por SQL.

Enterprise JavaBeans (EJB) 3.0

Los *Enterprise JavaBeans* (*EJP*, por sus siglas) son una de las API que forman parte del estándar de construcción de aplicaciones empresariales J2EE (ahora JEE) de Oracle *Corporation*. Permite realizar la administración automática de transacciones, seguridad, escalabilidad, concurrencia, distribución, acceso a ambientes portables y persistencia de datos. Su objetivo es simplificar el desarrollo de aplicaciones Java y estandarizar el API de persistencia para la plataforma Java (Monson-Haefel , 2004).

Jboss SEAM 2.1.1

Jboss SEAM es un framework desarrollado por Jboss que tiene como objetivo facilitar el desarrollo de aplicaciones sobre todo aplicaciones web. Es una plataforma de desarrollo de código abierto. SEAM integra tecnologías como *JavaScript* asíncrono y XML (AJAX), *JSF*, *JPA*, *EJB* y *Business Process Management* (BPM) (Liu, 2011). Su función principal radica en eliminar la capa artificial que existe entre EJB 3.0 y JSF y provee un sistema de anotaciones para integrar estos dos *frameworks*. Comparada con aplicaciones desarrolladas en otros *frameworks*, las aplicaciones SEAM son conceptualmente simples y requieren significativamente menos código (en Java y en XML) para obtener las mismas funcionalidades.

Java Runtime Environment (JRE)

JRE es el acrónimo de *Java Runtime Environment* y entendámoslo como conjunto de utilidades que permite la ejecución de programas Java sobre todas las plataformas soportadas. La Máquina Virtual Java (JVM, por sus siglas en inglés) es una instancia de JRE en tiempo de ejecución. Este interpreta el código Java y está compuesto además por las librerías de clases estándar que implementan el API de Java, es decir, actúa como un "intermediario" entre el sistema operativo y Java.

Jboss Server 4.2.2

Jboss es un servidor de aplicaciones Java EE de código abierto implementado en Java puro y el más utilizado actualmente en el mercado. Puede ser utilizado en cualquier sistema operativo donde se encuentre la JVM. El proyecto fue adquirido por la empresa *Red Hat* y se nutre de una red mundial de colaboradores. Jboss implementa todo el paquete de servicios de J2EE y provee servicios extendidos de almacenamiento de datos en memoria y de manera persistente (Redhat, 2010). Permite la integración de todas las tecnologías y herramientas utilizadas por *Jboss SEAM*.

Facelets 1.1

JavaServer Facelets es un *framework* para plantillas centrado en la tecnología JSF, por lo cual se integran de manera muy fácil. *Facelets* llena el vacío entre JSP (*JavaServer Pages*) y JSF, siendo una tecnología centrada en crear árboles de componentes y estar relacionado con el complejo ciclo de vida JSF (Liu, 2011).

Richfaces 3.3.1

Richfaces es una biblioteca de código abierto o lo que es lo mismo, librería de componentes web enriquecidos de código abierto y basada en el estándar JSF. Está desarrollado y mantenido por

Jboss. Se integra completamente dentro del ciclo de vida JSF. Permite crear interfaces de usuario modernas de manera eficiente y rápida, basadas en componentes listos para usar, altamente configurables en cuanto a temas y esquemas de colores predefinidos por el propio *framework* o desarrollados a conveniencia, lo que mejora grandemente la experiencia de usuario.

JGraphT 0.9.0 y JGraph 3.3.0

JGraphT y su JGraph biblioteca relacionada, proporciona soporte para gráficos matemáticos. Se puede modelar gráficos con bordes ponderados o no ponderados, también es compatible con grafos dirigidos. Puede crear gráficos basados en cualquier objeto Java incluyendo cadenas y archivos XML. JGraphT contiene el código para las estructuras de datos y algoritmos de modelado, mientras JGraph maneja la renderización y edición gráfica a través de una interfaz gráfica de usuario («JGraphT y JGraph», 2015). Es utilizada en la personalización porque Java como lenguaje de programación no tiene soporte nativo para la representación gráfica, por ende es necesario hacer usos de bibliotecas libres como JGraph para graficar valores de datos.

1.10. Conclusiones del capítulo

- ✓ El estudio de la Minería de Procesos demuestra que constituye una herramienta de mejora para los centros de salud e instituciones hospitalarias, éstas tienden a adoptar una visión de sus procesos por lo que es indispensable un correcto modelado e implementación de los mismos. Los análisis de procesos actuales están más centrados en cómo debería hacerse que en cómo se hace realmente por lo que dichos modelos en ocasiones se generan pareciéndose poco o nada a como realmente se está desarrollando la actividad y la Minería de Procesos surge para proveer la solución a dicho problema.
- ✓ La técnica *Fuzzy Miner* se ha puesto en práctica en el sector de la salud debido a las gráficas, modelos y secuencias que genera se han podido observar los pasos seguidos por los pacientes, conocer las excepciones e identificar cuellos de botella en los procesos lo que permite una vez obtenida toda la información ayudar a la toma de decisiones.

CAPÍTULO 2. FUZZY MINER: ANÁLISIS Y PROPUESTA DE SOLUCIÓN

En el presente capítulo se describen los patrones de diseño y arquitectura a utilizar, se analiza el *plugin Fuzzy Miner* de ProM, se presenta la propuesta de solución y se describe el procedimiento a seguir para la personalización del *plugin* sobre el HIS del CESIM que permitirá detectar la presencia de ruido en sus registros de eventos.

2.1. Propuesta de solución

El HIS como sistema de información posee un número considerable de datos alojados en su interior dentro de los que se encuentran los datos de eventos, los que no son aprovechados al no ser convertidos en información útil para la toma de decisiones, aunque sobre dicho sistema se desarrolló un componente para la extracción y transformación de las trazas de ejecución ésta información no es procesada para su posterior análisis.

Por otro lado, está la necesidad de ser cada vez más eficientes, lo que implica una búsqueda constante por mejorar los procesos. En este sentido, es conveniente usar Minería de Procesos para analizar los procesos del sistema y hacer un buen uso de la información almacenada. Además la utilización de esta disciplina puede ser de gran ayuda al permitir transparentar y analizar los procesos que les apoyaría en la toma de decisiones y así obtener procesos más eficientes.

Como se muestra en la **Figura 5** que representa el flujo de información de la propuesta de solución, el objetivo es integrar al HIS una personalización del *plugin Fuzzy Miner* para visualizar los modelos. El mismo analizará los registros de eventos generados por el componente que tiene el sistema para la extracción y transformación de las trazas. El fin de la propuesta de solución es mejorar el entendimiento y usabilidad de los modelos para los no expertos en el área de Minería de Procesos para ello descrito en lenguaje natural, se esconderá el algoritmo *Fuzzy Miner* detrás de una interfaz de usuario amigable que solo define parámetros comprensibles que sugieran tipos de análisis apropiados.

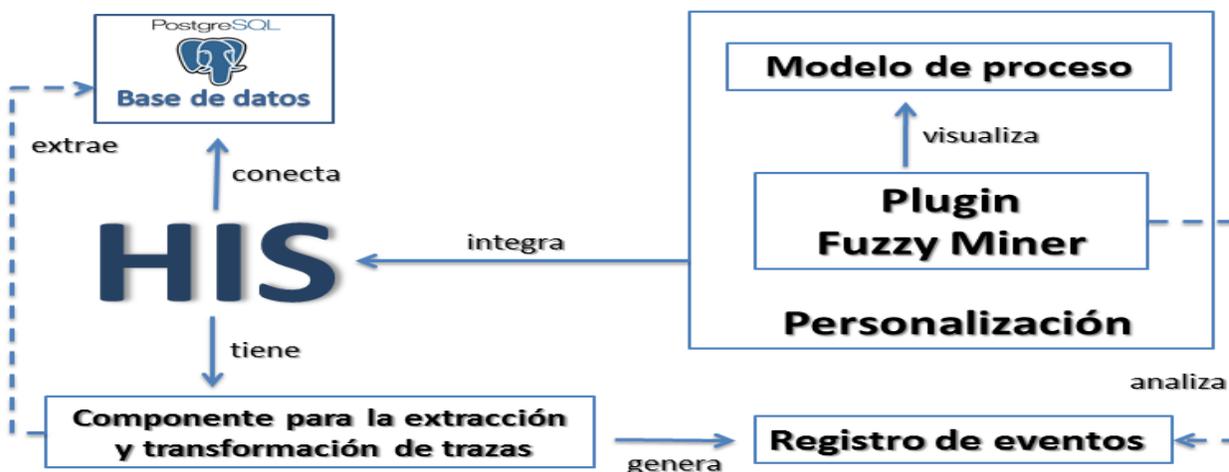


Figura 5: Propuesta de la solución.

Fuente: Elaboración propia.

2.2. Patrones de diseño y arquitectura

Los patrones de diseño expresan esquemas para definir estructuras de diseño (o sus relaciones) con las que se construyen sistemas de *software*. Los patrones de arquitectura expresan un esquema organizativo estructural fundamental para sistemas de *software*.

2.2.1. Patrón arquitectónico Modelo-Vista-Controlador

El HIS está desarrollado sobre la arquitectura Modelo Vista Controlador (MVC) un patrón arquitectónico que tiene como principal característica separar los datos de una aplicación en tres módulos identificables y con funcionalidades bien definidas: el Modelo, las Vistas y el Controlador, permitiendo flexibilidad y facilidad a la hora de hacer futuros cambios. (Burbeck, 2012).

El **modelo** es un conjunto de clases que representan la información del mundo real que el sistema debe procesar (Deacon, 2012). En el mismo se encuentran todas las funciones para consultar, insertar y actualizar información de la base de datos.

Las **vistas** son el conjunto de clases que se encargan de presentar al usuario la información contenida en el modelo. Una vista está asociada a un modelo del que pueden existir varias vistas asociadas al mismo. Estas pueden ser una página web o una parte de una página y obtiene del modelo solamente la información que necesita para desplegar y se actualiza cada vez que el modelo del dominio cambia por medio de notificaciones. (Deacon, 2012)

El **controlador** es un objeto que se encarga de dirigir el flujo del control de la aplicación debido a mensajes externos como datos introducidos por el usuario u opciones del menú seleccionadas por

él. A partir de estos mensajes, el controlador se encarga de modificar el modelo o de abrir y cerrar vistas. El controlador tiene acceso al modelo y a las vistas, pero las vistas y el modelo no conocen de la existencia del controlador (Burbeck, 2012).

En la **Figura 6** se muestra la interacción MVC de la personalización del *plugin Fuzzy Miner* para el HIS donde la vista envía una solicitud HTTP⁸ al controlador el cual construye el modelo difuso a partir de los datos que recibe del modelo y se devuelve a la vista donde es mostrado al usuario.

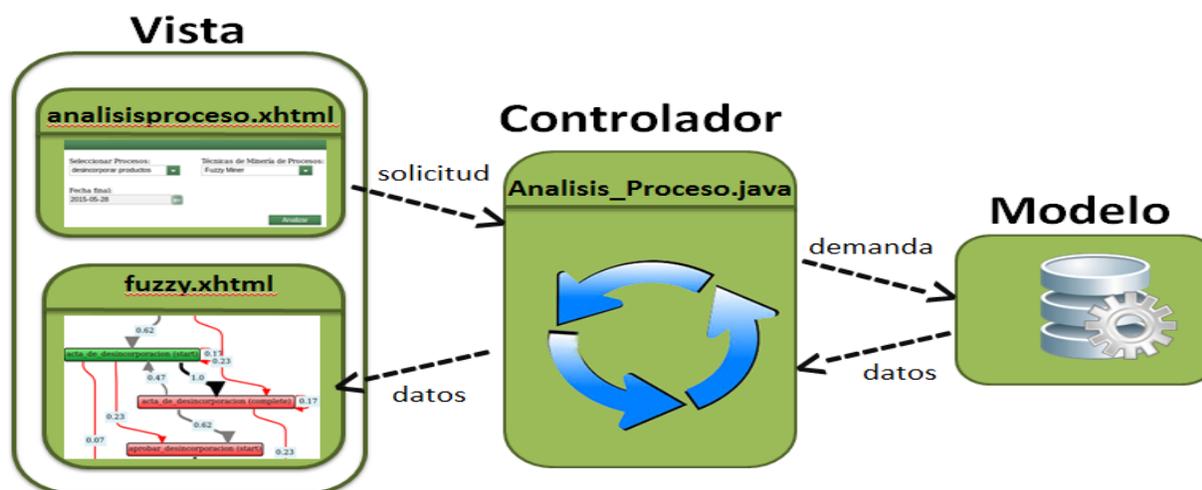


Figura 6: Interacción MVC.

Fuente: Elaboración propia.

2.2.2. Patrón de diseño GRASP

Los patrones GRASP representan los principios básicos de la asignación de responsabilidades a objetos expresados en forma de patrones. GRASP es el acrónimo para *General Responsibility Assignment Software Patterns* (Patrones Generales de Software para Asignar Responsabilidades) (EcuRed, 2015). Para el desarrollo de la solución se utilizan estos patrones, entre ellos el patrón experto, alta cohesión y el controlador con el objetivo de describir los principios fundamentales de diseño de objetos para la asignación de responsabilidades.

Patrones básicos de asignación de responsabilidades

Experto: Se encarga de asignar una responsabilidad al experto en información, o sea, aquella clase que cuenta con la información necesaria para cumplir la responsabilidad. Indica que la responsabilidad de la creación de un objeto o la implementación de un método debe recaer sobre

⁸HTTP: Abreviatura de la forma inglesa *Hypertext Transfer Protocol*, protocolo de transferencia de hipertextos, que se utiliza en algunas direcciones de internet.

la clase que conoce toda la información necesaria para crearlo. Se evidencia en la clase `FuzzyMinerPlugin.java` que tiene la responsabilidad de crear una instancia de la técnica *Fuzzy Miner* y permite crear los repositorios de métricas a partir de los parámetros de entrada y del registro de evento.

Alta Cohesión: Es una medida de cuán relacionados y enfocados están las responsabilidades de una clase. Una alta cohesión caracteriza a las clases con responsabilidades estrechamente relacionadas para que no realice un trabajo enorme. Se evidencia en las clases `FuzzyMinerPlugin.java` y `MetricsRepository.java` de tal manera que la clase `MetricsRepository.java` solamente obtiene los valores a partir de la clase `FuzzyMinerPlugin.java`.

Controlador: El patrón controlador es un patrón que sirve como intermediario entre una determinada interfaz y el algoritmo que la implementa, de tal forma que es la que recibe los datos del usuario y la que los envía a las distintas clases según el método llamado. Se recomienda dividir los eventos del sistema en el mayor número de controladores para poder aumentar la cohesión y disminuir el acoplamiento. Este patrón se evidencia en la clase `Analisis_Proceso.java`.

2.3. Análisis del plugin Fuzzy Miner de ProM

El *plugin Fuzzy Miner* fue desarrollado con el fin de optimizar y mejorar los procesos a los que se les aplique. Hace uso de técnicas para la extracción y agrupación de actividades con el objetivo de representar un proceso que sea comprensible por los analistas de procesos. El *plugin Fuzzy Miner* de ProM cuenta con una serie de métricas configurables derivadas de las mediciones del mismo para generar los modelos. A continuación se resumen las configuraciones presentadas en tres artículos en particular (Process Mining Group, 2009; Li, 2010; Van der Aalst, W.M.P. y Günther, 2007).

2.3.1. Configuración de métricas

Existen tres tipos de mediciones (significado unario, binario y correlación) y cada una tiene métricas de configuración:

Métricas de Significado Unario: Comportamiento de las actividades en el registro de eventos.



Figura 7: Interfaz de ProM para visualizar las métricas unarias.

Fuente: Elaboración propia.

- **Significado de frecuencia:** Esta dado por la cantidad de veces que se repite una actividad con respecto a todas las demás en un registro de eventos. La métrica es normalizada, por tanto, la actividad que más se repite toma el valor de uno mientras que el valor de frecuencia de las demás es calculado a partir de ella.
- **Significado de enrutamiento:** Esta dado por el balance que exista entre los arcos que entran a un nodo y los que salen de él. Mientras mayor sean las conexiones con otras actividades (mayor cantidad de arcos que entran y salen) será mayor significado de enrutamiento.

Métricas de Significado Binario: Comportamiento de las relaciones de precedencia (o aristas) entre nodos.

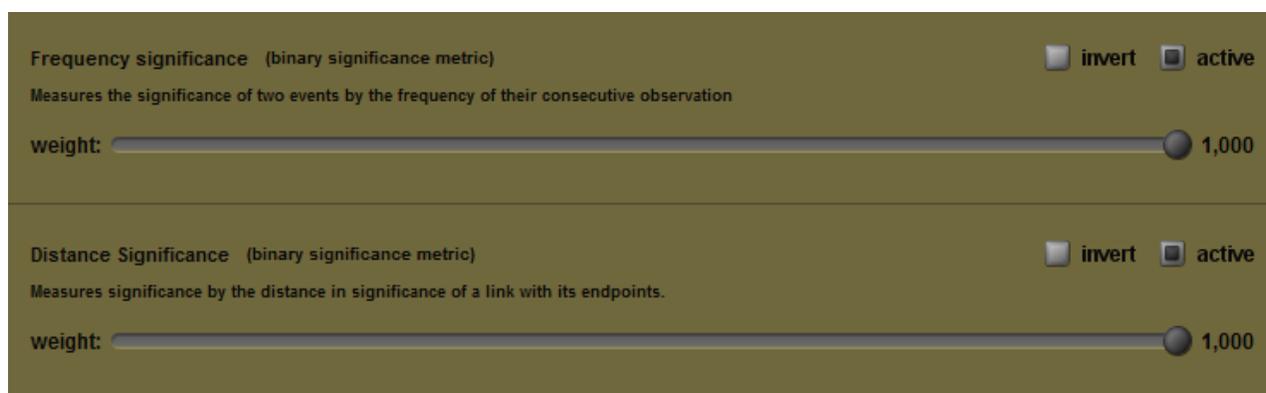


Figura 8: Interfaz de ProM para visualizar las métricas binarias.

Fuente: Elaboración propia.

Al igual que en el significado unario, la métrica de frecuencia es la más importante en el significado binario.

- Significado de distancia: Esta dado por la relación que exista entre el significado de la actividad origen con el significado de la actividad objetivo. El valor que tome será el menor valor de significado entre las dos actividades. Esta métrica es indispensable para aislar comportamientos de interés.

Métricas de Correlación Binaria: Mide que tan relacionado está una actividad de otra. La correlación binaria es la que maneja la decisión entre la agregación o la abstracción de los comportamientos menos significativos.

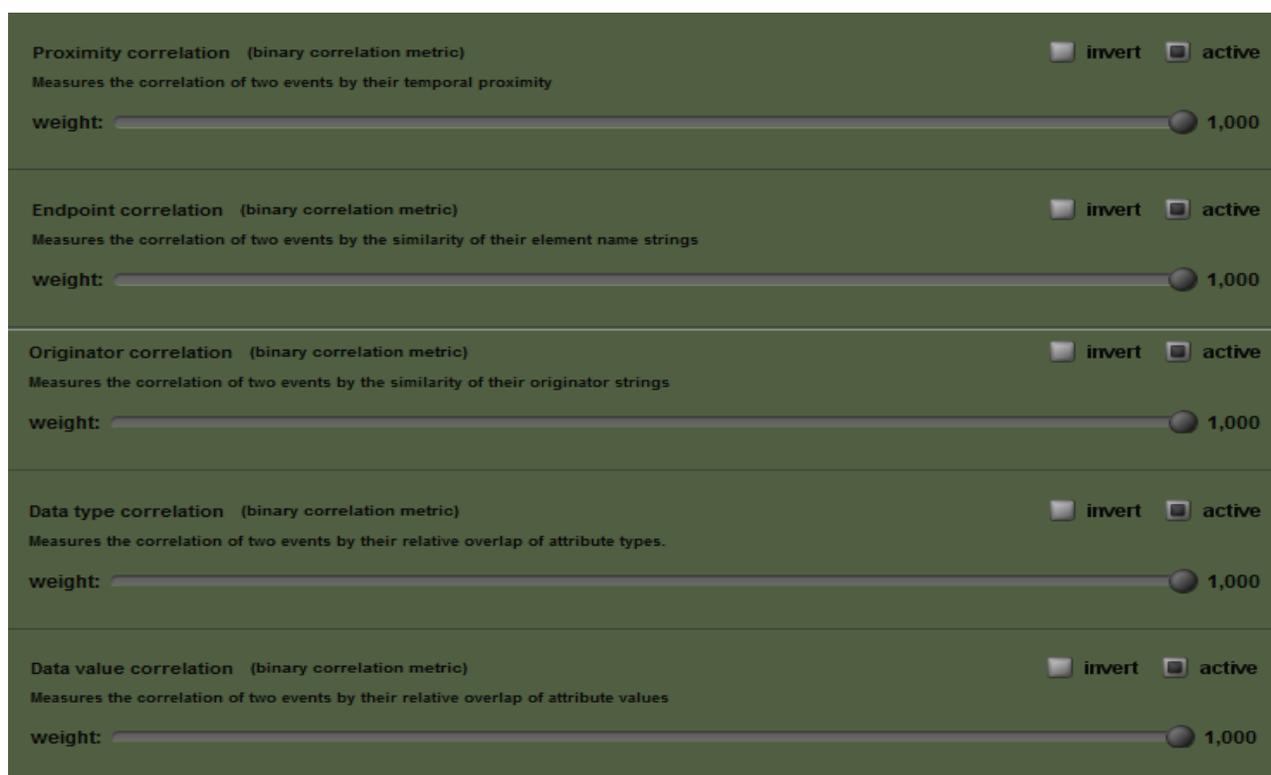


Figura 9: Interfaz de ProM para visualizar las métricas de correlación binaria.

Fuente: Elaboración propia.

- La correlación de proximidad: Evalúa la ocurrencia de un evento con respecto a otro, mientras menor tiempo de ejecución mayor correlación. Esta métrica es importante para identificar *cluster* de eventos que corresponde a una sola actividad lógica.
- La correlación originadora: Se tiene en cuenta el nombre de la persona que llevo a cabo la sub-secuencia de dos eventos. Mientras mayor similitud tengan los nombres mayor será el valor de correlación.
- La correlación de punto final: Es similar a la correlación originadora sin embargo la comparación que se tiene en cuenta es el nombre del recurso con el nombre de la actividad entre

las dos sub-secuencia de actividades. Mientras más similares sean los nombres mayor será el valor de correlación.

- La correlación tipo de datos: En la mayoría de los registros, se incluyen atributos adicionales. Estos atributos son interpretados en el contexto en el que son usados. Esta métrica evalúa cada actividad. Mientras mayor cantidad de atributos existan y sean comunes entre las dos actividades, mayor será su correlación.
- La correlación valor de datos: Es semejante a la correlación tipo de datos pero no tiene en cuenta el tipo de datos si no los valores que contengan los atributos que son comunes. Un pequeño cambio en un atributo comprometerá el valor de la correlación.

La configuración de métricas cuenta con una lista de parámetros configurables. Cada métrica tiene el mismo conjunto de opciones de configuración que le ayudan a optimizar las medidas respecto a su situación específica:

Peso: Todas las métricas (significado unario, binario y correlación) tendrán un valor específico cuando se realice la minería. Por ejemplo, para enfatizar una métrica específica, se reduce el peso de todas las demás métricas.

Invertir: Si esta casilla de verificación está activada, todas las mediciones de las métricas se invierten lo que significa que el nuevo valor será $(1 - \text{valor_original})$. Todos los valores recogidos por una métrica se normalizaran, tal que la medición más alta será igual a 1.0. Esto puede ser una herramienta muy útil si, por ejemplo, se quieren que los eventos altamente frecuentes sean considerados menos importantes.

Activo: Utiliza esta opción cuando se cree que una métrica específica no contribuye a mejores resultados, o sea incluso contra-productivo. Tenga en cuenta que al establecer una métrica en inactivo no mejorará el rendimiento; el *plugin Fuzzy Miner* está optimizado para ejecutarse en cualquier personalización de los ajustes.

2.3.2. Configuración de medición

El *Fuzzy Miner* no se limita a medir solamente las medidas de significación y correlación de dos actividades sino que también puede medir las relaciones a largo plazo.

Configuración del punto de medida

En el área de configuración del punto de medida cuenta con un histograma que visualiza el número de puntos de medición por eventos (por ejemplo, el número de barras de histograma) y su

factor de evaluación (altura de las barras). La **Maximal event distance** (distancia máxima de eventos) define el número de puntos de medida.

Atenuación

La atenuación más simple es la **Linear attenuation** (atenuación lineal), que garantizará la atenuación con la distancia de eventos. La atenuación **Nth root whit radical** (raíz n-ésima con radicales) permite la atenuación negativa exponencial para configurar la función *Nth root*. Un valor relativamente alto atenuará progresivamente los puntos de mayor distancia de medida, que es útil cuando desea centrarse en las relaciones a corto plazo. Un valor relativamente bajo ayuda a identificar mejor las actividades que constituyen ruido en el registro de eventos (Process Mining Group, 2009).

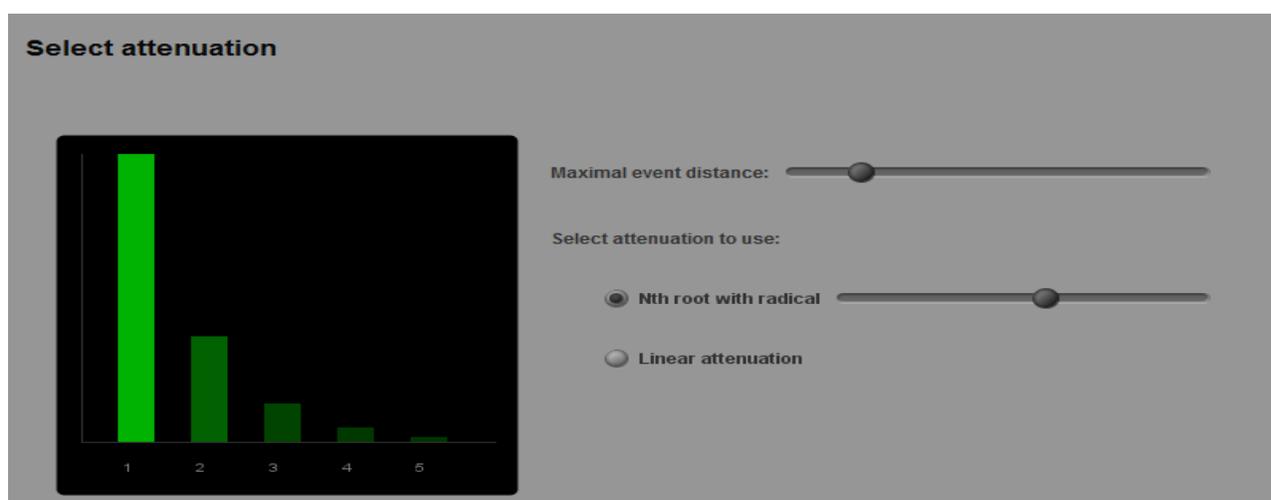


Figura 10: Interfaz de ProM para visualizar la distancia máxima de evento y la atenuación.

Fuente: Elaboración propia.

2.3.3. Proceso de minería

Las configuraciones son ajustables, una vez confirmadas se pone en marcha el proceso de minería y cuando concluya se obtiene un modelo difuso como se muestra en la **Figura 11**.

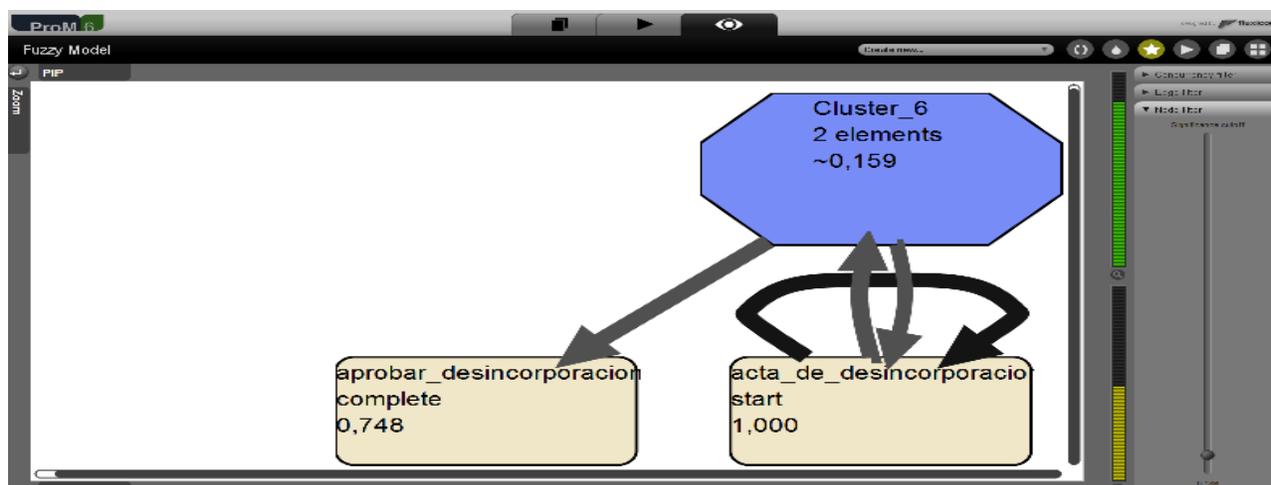


Figura 11: Interfaz de ProM de un modelo difuso.

Fuente: Herramienta ProM.

El objetivo final de *Fuzzy Miner* es crear una representación gráfica apropiada del proceso que se encuentra en el registro. La notación gráfica es bastante sencilla, los nodos cuadrados representan actividades. Los comportamientos correlacionados menos significativos se descartan, es decir, los nodos y arcos que entran en esta categoría se eliminan de la gráfica. Mientras que los grupos coherentes de comportamientos menos significativos y altamente correlacionados se representan de forma agregada en un *cluster*, quienes a su vez son representados como octágonos azules. Los componentes internos de los *cluster* y su estructura pueden ser explorados dando clic sobre él.

Por otra parte, se cuenta con diferentes filtros para la modelación, el **Filtro de nodo** controla la cantidad de actividades que se incluirán en el gráfico. El **Filtro de Aristas** (dos tipos, mejor filtro de aristas y filtro de aristas difusos) influye en la cantidad de aristas y su selección que se incluirán en el gráfico y el **Filtro de concurrencia** se utiliza para resolver conflictos entre dos clases de evento.

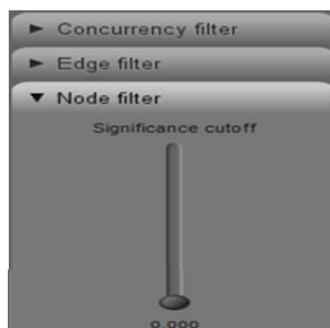


Figura 12: Ejemplo de filtros del plugin Fuzzy Miner de ProM.

Fuente: Elaboración propia.

Además, el plugin cuenta con las Vistas de métricas unarias y Vistas de métricas binarias.

Vista de métricas unarias

La **Figura 13** muestra las mediciones de todas las métricas para todas las clases de eventos en un gráfico bidimensional de curva múltiples. Esto le permite obtener una retroalimentación rápida e intuitiva sobre si es apropiada la configuración de métricas. De pasar el cursor sobre el gráfico se revelarán los valores de medición precisos de cada clase de evento. En las casillas coloreadas en la parte inferior izquierda de la imagen se cambia la visibilidad de cada métrica en el gráfico.

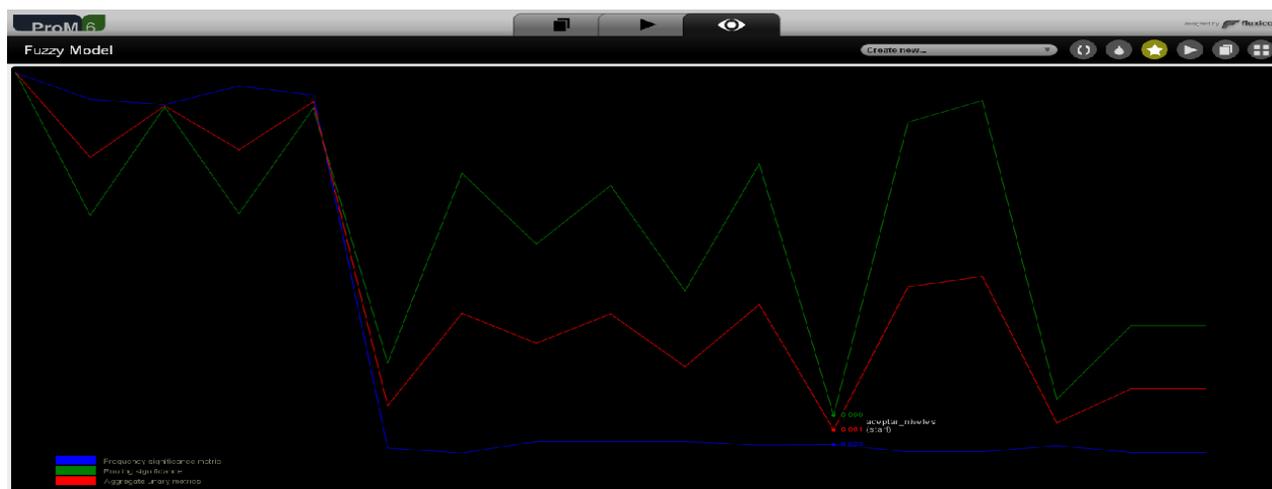


Figura 13: Interfaz de ProM de la Vista de métricas unarias.

Fuente: Elaboración propia.

Vista de métricas binarias

Las métricas binarias se muestran en forma de una matriz de color (**Figura 14**). Si se sitúa el cursor sobre la matriz se revelarán los valores de medición precisos para cada combinación de clase de evento.

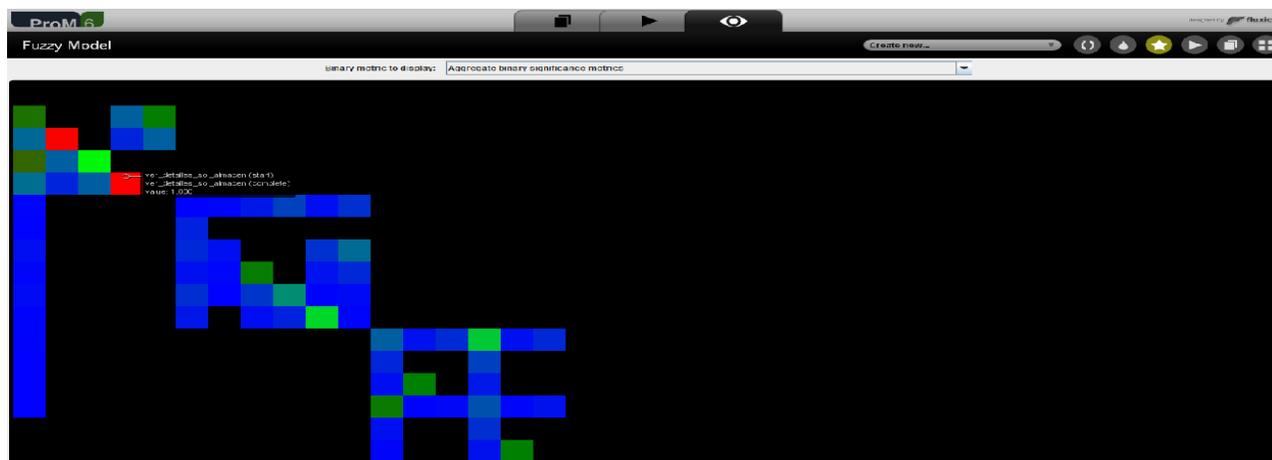


Figura 14: Interfaz de ProM de la Vista de métricas binarias.

Fuente: Elaboración propia.

2.4. Procedimiento para la detección de ruido con el plugin Fuzzy Miner para el HIS

El procedimiento para la detección de ruido aplicando el *plugin Fuzzy Miner* de ProM tiene como objetivo proveer al proceso de desarrollo de *software* del HIS de un *plugin* que contribuye a la visualización de un modelo que sea entendible con la abstracción de los detalles no deseados. Para su aplicación deben cumplirse las premisas establecidas para el componente de extracción de registros de eventos del HIS, por ejemplo, el *log* de eventos debe estar en formato *.XES⁹ que registre como información mínima asociada al evento un identificador, el originador y la marca de tiempo. El procedimiento propone tres fases para su aplicación: “Determinar las configuraciones necesarias para el HIS”, “Integrar el *plugin Fuzzy Miner* al HIS” y “Mostrar el modelo desde el HIS”.

Para mayor comprensión de la propuesta de solución se generaron los Productos de Trabajo correspondientes al requisito funcional “Generar modelo de proceso” del componente para la detección de variabilidad en procesos hospitalarios, disponibles en **Anexo 2**, **Anexo 3** y **Anexo 4**.

2.4.1. Fase 1: Determinar las configuraciones necesarias para el HIS

Cuando se analizan los registros de eventos que genera el Sistema de Información Hospitalaria se hace evidente que dichos *logs* contienen como información asociada al evento a un identificador, el originador y la marca de tiempo (*name*, *resource*, *timestamp*; por sus nombres en inglés respectivamente), entre otros datos.

A continuación se muestra un ejemplo de un registro de eventos del proceso “Desincorporar productos” del módulo almacén del HIS, donde se evidencia que las premisas planteadas para aplicar la técnica en cuestión, se cumplen.

<trace>

1 <string key="concept:name" value="17438"/>

2 <string key="description" value="Simulated process instance"/>

<event>

3 <string key="org:resource" value="mirandae@pdvsa.com"/>

4 <date key="time:timestamp" value="2012-06-20T15:47:03.703.000+01:00"/>

5 <string key="concept:name" value="acta_de_desincorporacion"/>

⁹.XES: es un estándar XML para los registros de eventos. El estándar ha sido adoptado por la IEEE Task Force on Process Mining como el formato de intercambio de registros de eventos por defecto.

```

        6      <string key="lifecycle:transition" value="start"/>
      </event>
</trace>

```

Línea 1 concept:name; muestra el identificador de la instancia del proceso.

Línea 2 description; muestra la descripción del proceso.

Línea 3 org:resource; muestra quien fue el que ejecuto la actividad.

Línea 4 time:timestamp; muestra la fecha y hora en que se ejecutó la actividad.

Línea 5 concept:name; muestra lo que se pretende hacer.

Línea 6 lifecycle:transition; muestra el ciclo de vida de la transición, en este caso puede tomar dos valores (*start*, *complete*).

Medidas y métricas

Según una investigación de Chamorro y Maturana (Chamorro y Maturana, 2013) como algunas herramientas (*plugin*) que tiene ProM permiten ser configuradas por ejemplo *Fuzzy Miner*, recomiendan que de no conocerse cabalmente lo que implica cada una de las medidas y métricas, se usen las configuraciones por defecto, ya que el modelo generado puede sufrir notorias variaciones al realizar modificaciones (**Anexo 5**).

En la personalización del *plugin Fuzzy Miner* para HIS se necesita obtener toda la información posible a partir de los datos que se guardan, por tanto se usan todas las métricas con sus valores por defecto con el objetivo de generar todas las actividades y sus relaciones para ir descartando del modelo las que representan ruido según el criterio del usuario.

De los parámetros configurables de las métricas se personaliza el Peso y Activo. El Peso representará el valor que va a tener la métrica, su valor por defecto es uno. Con el Activo siempre se tiene presente el valor que contiene la métrica para generar el modelo. En las configuraciones de medición, el punto de medida tendrá un valor máximo de distancia de cuatro y la atenuación *Nth root whit radical* con un valor de 2.7^4 por ser los mejores valores según (Process Mining Group, 2009) para tratar el ruido en un registro de eventos.

Proceso de minería

A la hora de visualizar el modelo solo se muestran nodos primitivos, no se personaliza el *cluster* (actividades con baja frecuencia y alta correlación) por la necesidad de mostrar a los usuarios finales las actividades que tienen baja frecuencia. Ruido es comportamiento excepcional (Van der Aalst, W.M.P., 2011a) y según Aalst este tipo de comportamiento es denominado de baja frecuencia que difiere de la mayoría de las trazas.

Es válido mencionar que la definición de ruido puede ser un poco ilógica. El término "ruido" se utiliza a veces para referirse a eventos incorrectos, es decir, los errores que se registran a la hora de generar una traza, para esta investigación y como interpretación de "ruido" en el campo de Minería de Procesos se utiliza éste término para referirse al comportamiento raro y poco frecuente "atípicos" en lugar de errores relacionados con el registro de eventos.

Filtros

Los filtros originales no se personalizan, el usuario final para hacer uso de ellos necesitara tener conocimientos de lo que significa cada uno, sin embargo se crean dos filtros el "Frecuencia de actividades" y "Frecuencia de aristas". El primero permite mostrar las actividades dependiendo del porcentaje que se quiera visualizar y responde al valor de las métricas unarias. Mientras que el segundo permite mostrar las relaciones de precedencia de dos actividades dependiendo del porcentaje que se quiera visualizar y responde a las métricas binarias y la correlación.

Vistas

La técnica *Fuzzy Miner* además del modelo que genera con las relaciones de las actividades también genera otras vistas para ver cómo se comportan los valores, estas vistas representan las métricas unarias, binarias y las correlaciones. Las vistas no se personalizan para interpretar sus valores se necesita tener conocimiento previo de cómo se calcula cada una de las métricas y que significa cada valor, por tanto no es aconsejable mostrarlo solamente tenerlos en cuenta de forma que las configuraciones generen un modelo útil y comprensible.

2.4.2. Fase 2: Integrar el plugin Fuzzy Miner al HIS

En la **Figura 15** se muestra el diagrama de clases donde se encuentran las clases necesarias para integrar el componente desarrollado al HIS.

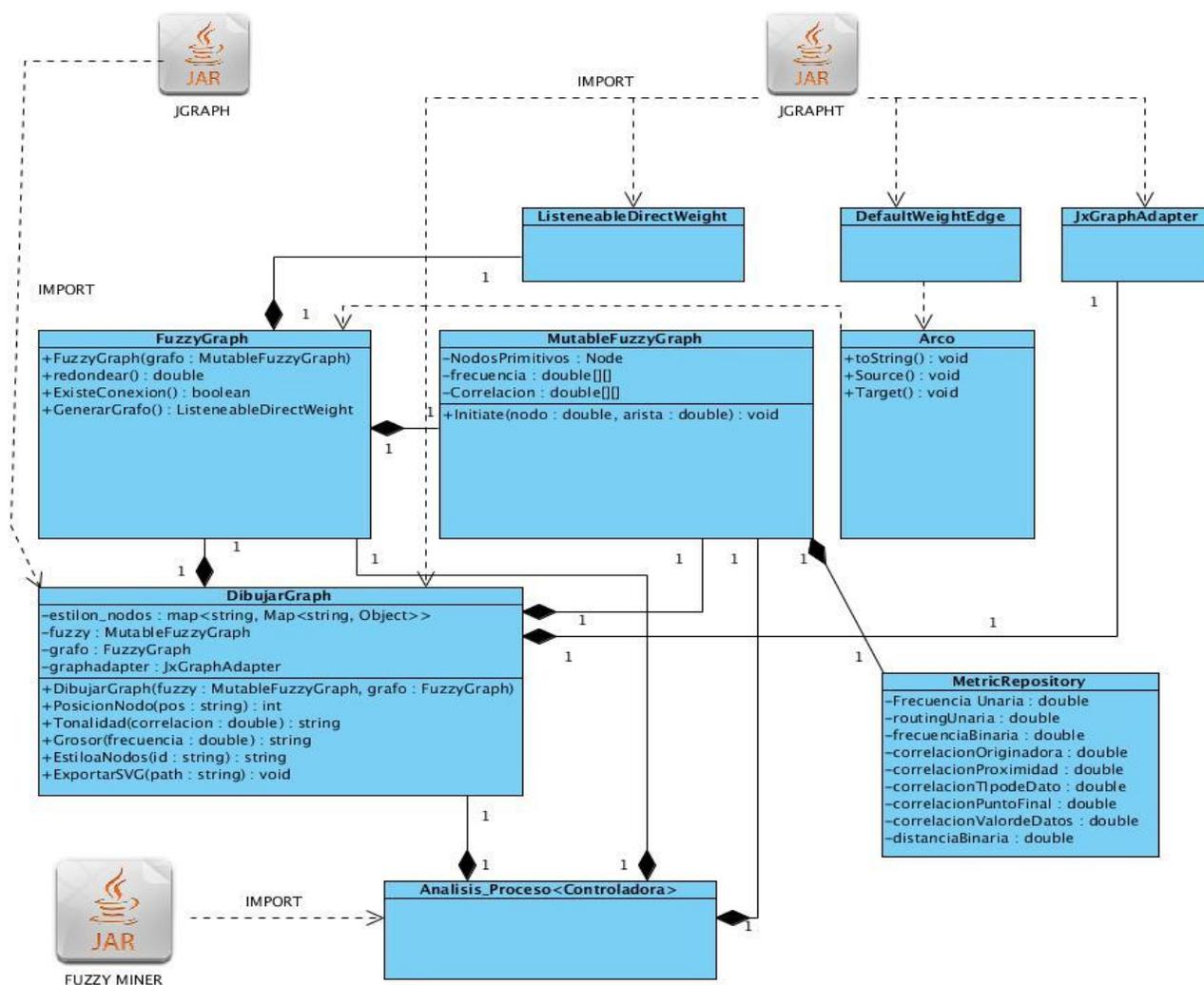


Figura 15: Diagrama de clases.

Fuente: Elaboración propia.

La clase FuzzyGraph.java crea el árbol con los valores de las métricas y DibujarGraph.java dibuja el modelo que será mostrado (Figura 16). La clase Analisis_Proceso.java es la controladora del componente para la detección de variabilidad en procesos hospitalarios, aquí se crean las instancias de las clases involucradas para integrar el componente y modelar los procesos (Figura 17).

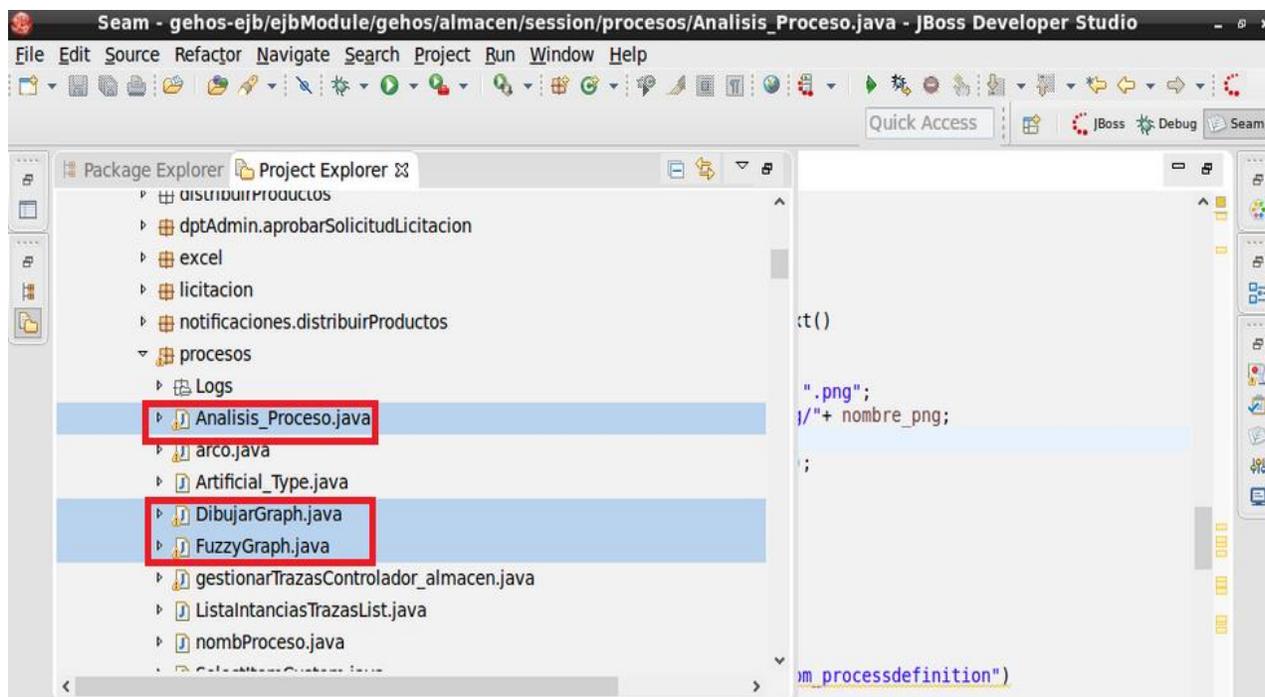


Figura 16: Interfaz de Jboss de la jerarquía de clases del componente.

Fuente: Elaboración propia.

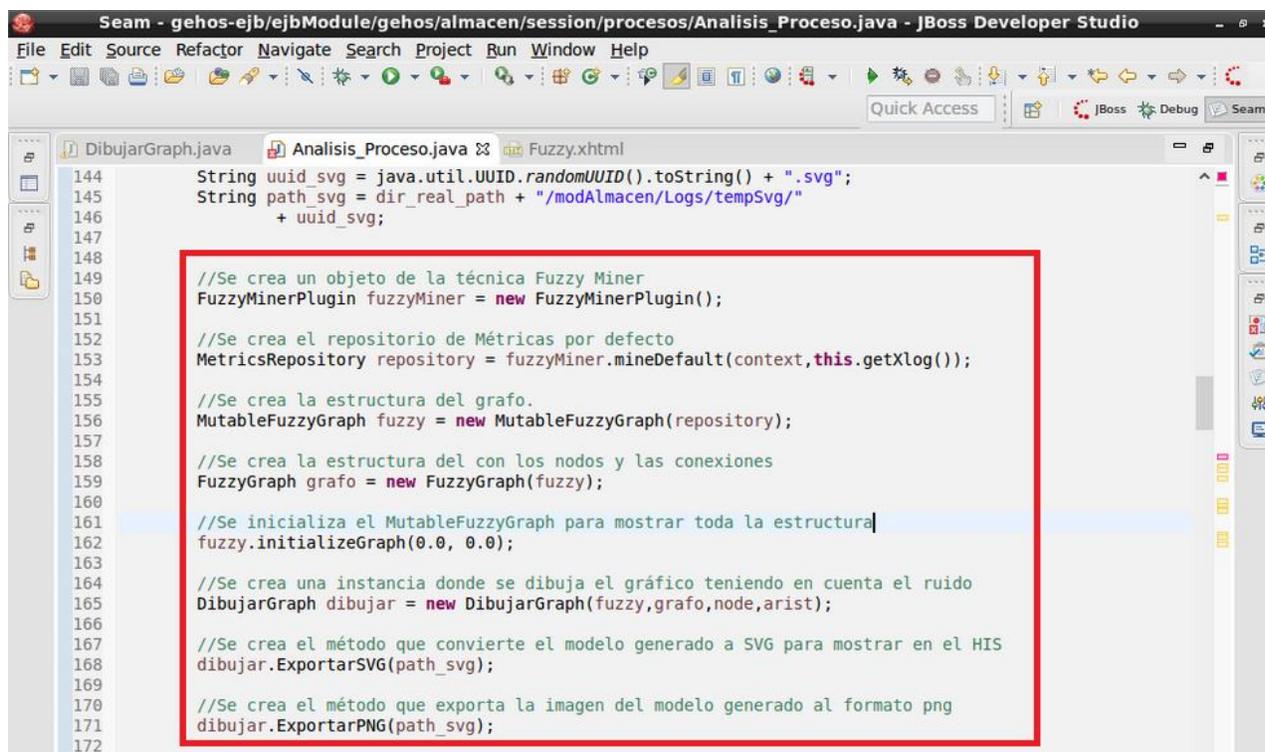


Figura 17: Interfaz de Jboss de las instancias de la clase Analisis_Proceso.java

Fuente: Elaboración propia.

2.4.3. Fase 3: Mostrar el modelo desde el HIS

Para mostrar el modelo desde el sistema solo es necesario seleccionar el proceso, el tipo de análisis que se desea y definir las fechas Desde – Hasta del registro de eventos que se va a generar (**Figura 18**). El modelo puede ser cancelado presionando la opción “Cancelar”.

The screenshot shows a web interface window titled "Parámetros de entrada". It contains four input fields arranged horizontally: "Procesos" with a dropdown menu showing "desincorporar productos", "Tipo de análisis" with a dropdown menu showing "Ruido", "Desde" with a date field showing "2010-06-01", and "Hasta" with a date field showing "2015-06-15". Below these fields are two buttons: "Generar" and "Cancelar".

Figura 18: Interfaz del HIS de los parámetros de configuración para el modelado.

Fuente: Elaboración propia.

La **Figura 19** muestra la interfaz de la personalización del técnica, al igual que el resto de las técnicas para el diagnóstico utiliza como entrada un registro de eventos y genera como salida un modelo difuso en el que se representan dos tipos de actividades y tres tipos de aristas con diferentes grosor y color que las identifican según la frecuencia y la correlación respectivamente, las actividades en verde significan que son actividades de alta frecuencia, es decir, que se repite un gran número de veces en el registro y las actividades de color rojo son las de baja frecuencia, se repiten pocas veces dentro del registro de eventos.

Las tonalidades definen los valores de correlación, las aristas de color negro significan que la relación tiene una alta correlación, mientras que el gris oscuro define a la correlación media y el gris claro las relaciones de poca correlación. El grosor de las aristas determina la frecuencia de las relaciones, mayor grosor, mediano grosor y poco grosor. Independiente al grosor y tonalidad, las relaciones que tengan un valor inferior al del filtro "Frecuencia de aristas", resaltarán en color rojo en el modelo. Los valores utilizados para diferenciar entre tonalidad y grosor, son los definidos en el *plugin Fuzzy Miner* de ProM (**Anexo 5**).

Ruido

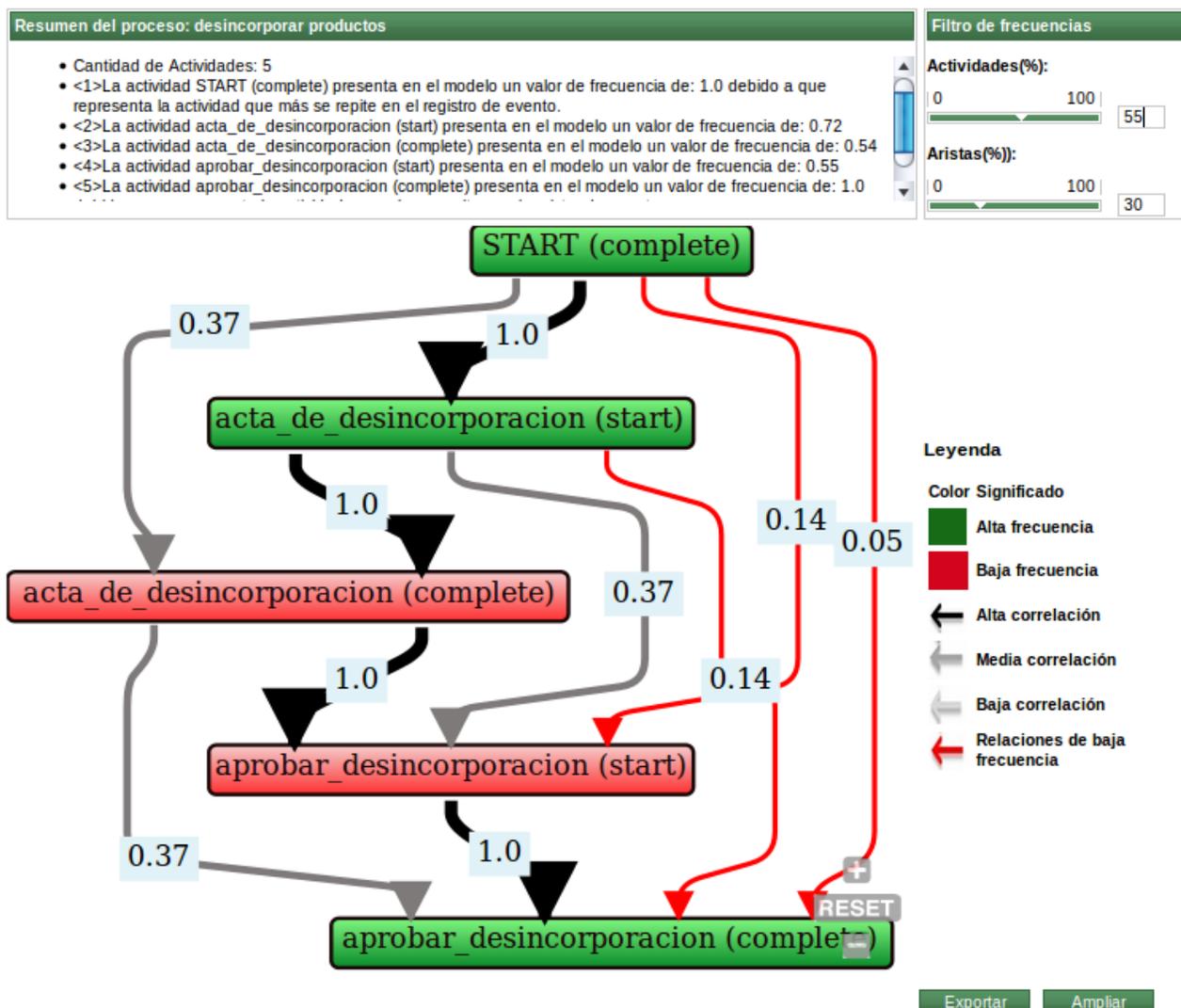


Figura 19: Ejemplo de la personalización de Fuzzy Miner en el HIS.

Fuente: Elaboración propia.

En la interfaz, la parte superior izquierda contiene un resumen del proceso que se esté analizando que muestra en lenguaje natural el comportamiento reflejado en las actividades del modelo obtenido. En la parte superior derecha se encuentran los filtros cuyos valores son ajustables para que el usuario adapte el modelo a sus necesidades y lo que realmente le interesa que sea mostrado en el momento que lo requiera.

En la parte central se muestra el modelo con la leyenda que provee una explicación de los elementos que intervienen en el mismo, los valores sobre los arcos representan la frecuencia de la relación de las actividades y en la parte inferior se cuenta con la opción de exportar el modelo a formato *.png para su posible análisis en otros entornos y conjuntamente posee la opción para

ampliar el modelo que permite aislarlo a una nueva pestaña del navegador, estas opciones forman parte de los requisitos del componente que genera el registro de eventos.

2.5. Modelo obtenido con Fuzzy Miner desde el HIS y comprensión del mismo

A continuación se muestra un ejemplo (**Figura 20**) de un modelo *Fuzzy Miner* que contiene los caminos o actividades más repetidas (secuencias principales) en el flujo de información del proceso “Desincorporar producto” del módulo Almacén. El modelo incluye las siguientes actividades en su flujo esencial: “START (complete)”, “acta_de_desincorporacion (start)”, “acta_de_desincorporacion (complete)”, “aprobar_desincorporacion (start)” y “aprobar_desincorporacion (complete)”.

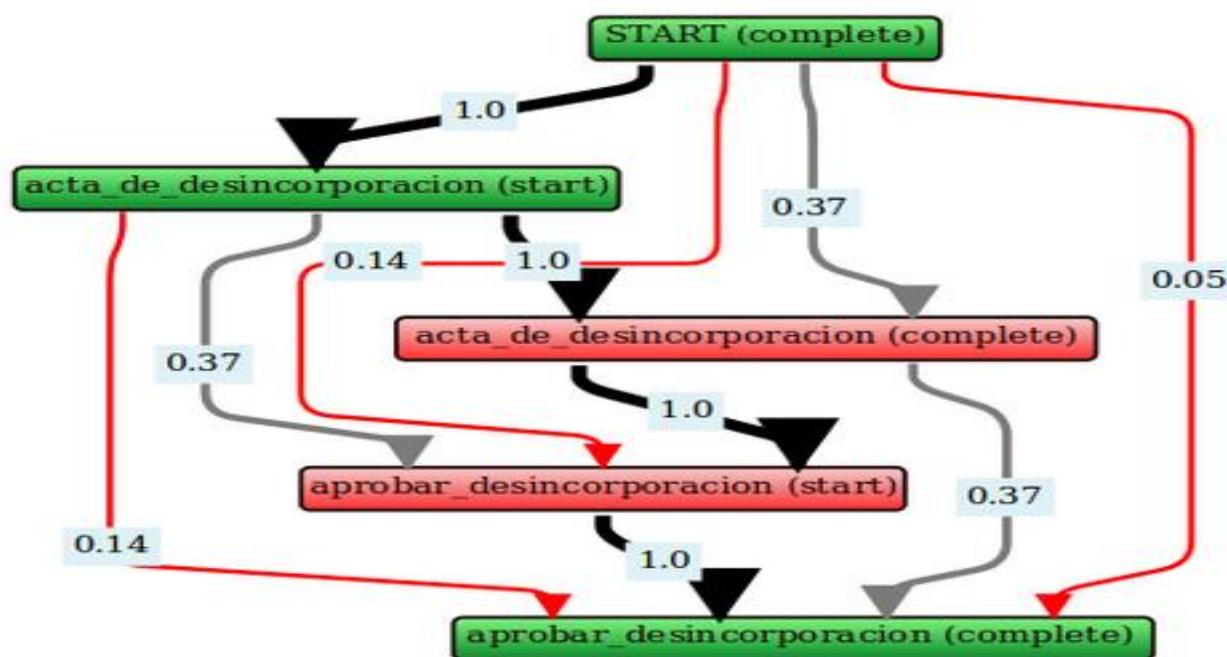


Figura 20: Secuencia principal del proceso con *Fuzzy Miner*.

Fuente: Elaboración propia.

En el modelo se evidencia consistencia de los datos almacenados. Por ejemplo, la actividad “aprobar_desincorporacion (complete)” está representada como una tarea final del proceso lo que demuestra la semejanza con la forma en que se realiza el proceso real. Cada nodo del modelo representado, muestra la significación que posee su correspondiente actividad. A partir de estos valores se identifica como la actividad más importante para la estructuración del modelo del proceso “START (complete)”. Se identifican, en este caso, como secuencias principales del proceso, inicio - “START (complete)”, “acta_de_desincorporacion (start)”, “acta_de_desincorporacion (complete)”, “aprobar_desincorporacion (start)” y fin -

“aprobar_desincorporacion (complete)”, se utiliza como criterio el color y el grosor de los arcos del modelo.

La comprensión del modelo del proceso (en términos de precisión y eficiencia) es una función de las características del modelo y de las características del usuario que interpreta el modelo. El diseño visual del modelo constituye un factor primordial para su análisis. Se han identificado factores que ayudan en el entendimiento del modelo de proceso tales como: notación de modelado, presentación visual y complejidad del modelo de proceso. Adicionalmente, se han analizado características del modelo que enriquecen su comprensión con respecto al *plugin Fuzzy Miner* de ProM tales como el resaltado de colores y el estilo gramatical de las etiquetas de texto.

2.6. Conclusiones del capítulo

- ✓ Con la utilización del algoritmo de descubrimiento *Fuzzy Miner* y a diferencia de otras técnicas, se ofrece una vista de alto nivel del modelo con la abstracción de los detalles no deseados, además permitió agrupar tareas donde cada tarea pertenece a un único nodo y fue empleada en el diagnóstico de registro de eventos, lo cual permitió realizar análisis preliminares al descubrimiento basados en la correlación entre las tareas y la importancia de una secuencia lo que brinda gran versatilidad al realizar análisis y tomar decisiones en cuanto a los resultados.
- ✓ El uso del *plugin Fuzzy Miner* permite con un mismo registro de eventos obtener varios mapas difusos diferentes dependiendo del uso previsto y los aspectos de interés.
- ✓ La personalización del *plugin Fuzzy Miner* de ProM se desarrolló basándose en conceptos para el Sistema de Información Hospitalaria del CESIM lo cual posibilita la construcción de modelos de acuerdo a las necesidades del personal médico y de forma rápida e intuitiva.

CAPÍTULO 3. VALIDACIÓN DE LA SOLUCIÓN PROPUESTA

En este capítulo se describen las validaciones realizadas a la personalización propuesta como complemento del Sistema de Información Hospitalaria. Haciendo uso del *plugin Fuzzy Miner* desde el HIS se evalúan un conjunto de procesos y se analizan los resultados obtenidos.

3.1. Validación

Las validaciones están basadas en la ejecución, revisión y retroalimentación de la personalización. Dicho de otro modo son pruebas específicas, concretas y exhaustivas para probar y validar la solución en cuanto a lo que hace, lo que debe y sobre todo lo que se ha especificado.

Con las validaciones realizadas se pretende:

- Comprobar la exactitud de interpretación del término ruido de la personalización del *plugin Fuzzy Miner* desde el HIS.
- Comprobar que el modelo difuso generado por el *plugin* desde el HIS contiene inicialmente todas las actividades del registro de eventos que se esté analizando.
- Comprobar la calidad del *Fuzzy Miner* desde el HIS.

3.1.1. Cantidad de actividades

Se utiliza el proceso “Solicitar productos” del módulo almacén del HIS que cuenta con 15 actividades diferentes y un total de 5713 ocurrencias. El registro de eventos recopila 1140 ejecuciones del proceso en las fechas comprendida entre 1 de mayo del 2011 y 4 de abril del 2015.

El objetivo que persigue la prueba es comprobar que todas las actividades del *log* son modeladas con la personalización del *plugin Fuzzy Miner* en el HIS. Haciendo uso del *plugin Add identities to log* de ProM (**Figura 21**) se sabe que el registro de eventos cuenta con 15 actividades del proceso “Solicitar productos” (**Figura 22**), así como la cantidad de instancias y otros detalles.

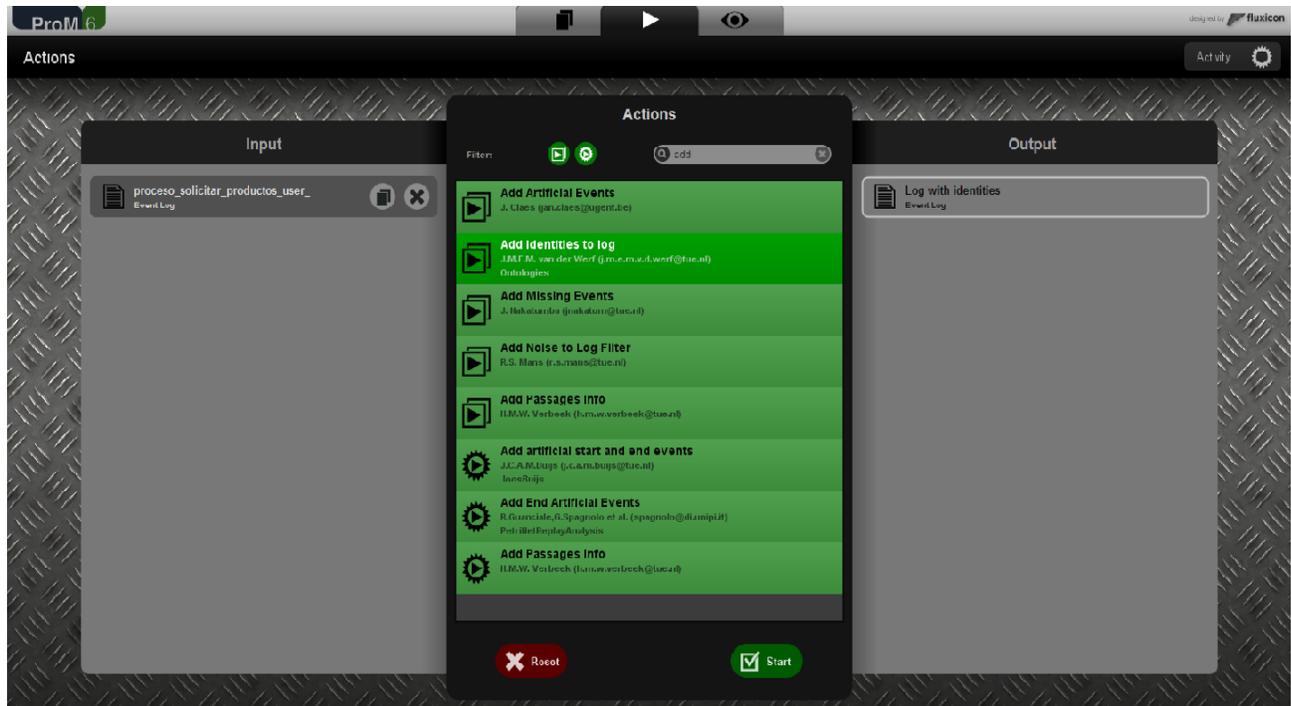


Figura 21: Interfaz de la herramienta ProM de selección del plugin Add identities to log.

Fuente: Elaboración propia.

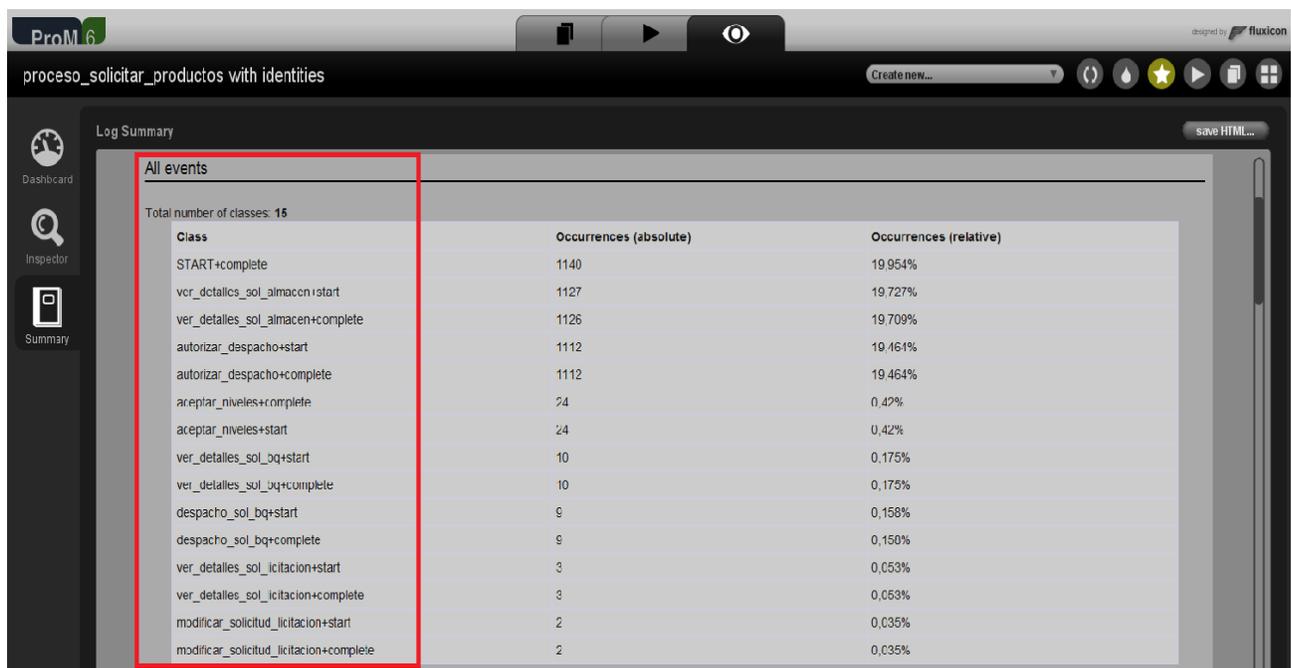


Figura 22: Interfaz de la herramienta ProM del plugin Add identities to log que muestra la cantidad de actividades.

Fuente: Elaboración propia.

Una vez generado el modelo difuso sobre el mismo registro de eventos del proceso “Solicitar productos” con la personalización del *plugin Fuzzy Miner* (Figura 23), se observan las 15 actividades con que cuenta el *log*.

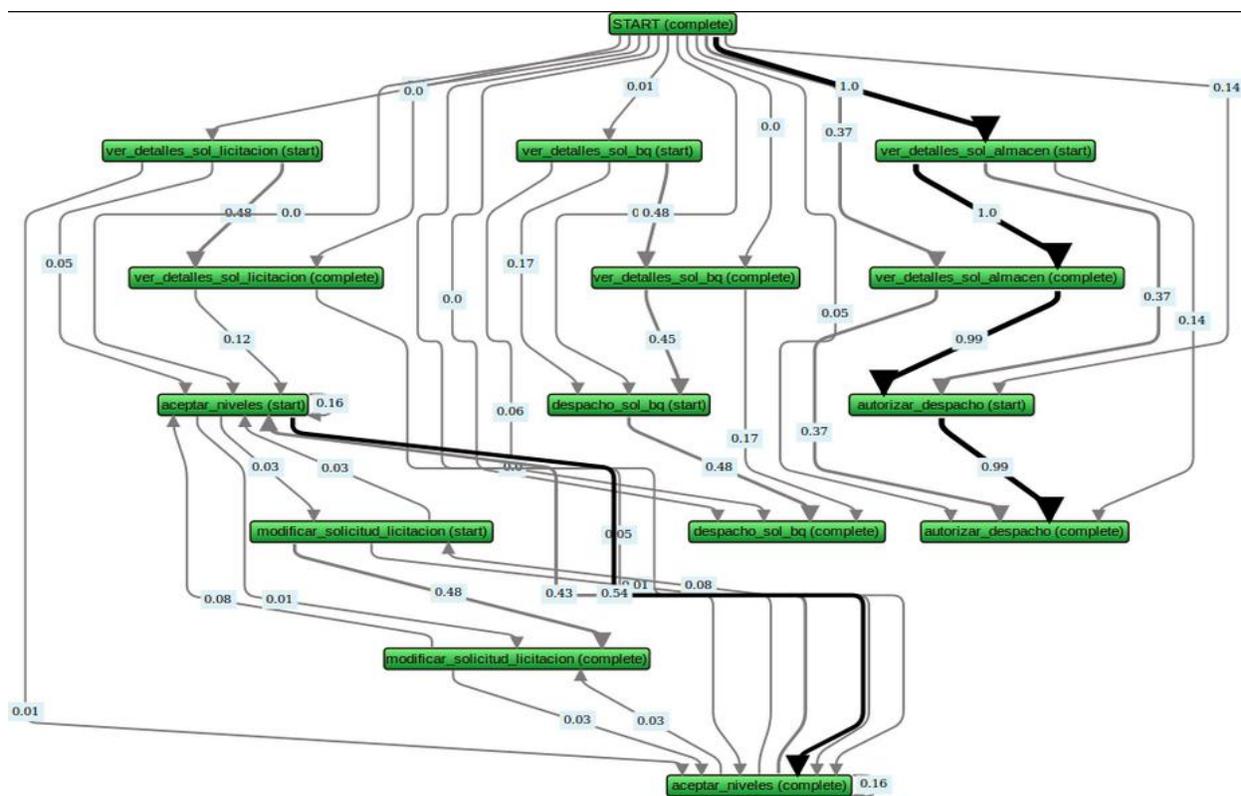


Figura 23: Interfaz del proceso “Solicitar productos” con el plugin Fuzzy Miner desde el HIS.

Fuente: Elaboración propia.

Como se muestra en el modelo obtenido la personalización de la técnica genera inicialmente todas las actividades del registro de eventos que se esté analizando.

3.1.2. Identificación de ruido en actividades

La validación tiene como objetivo comprobar que la solución desarrollada identifica de manera correcta las actividades que se consideran ruido según el usuario dentro de un registro de eventos. La validación tiene dos enfoques, el primero está orientado a comparar los modelos generados sobre un mismo registro de eventos tanto por la personalización como por la herramienta ProM y el segundo está dirigido a comprobar que para un valor de frecuencia establecido con anterioridad y un análisis manual de actividades deben ser visualizadas por el componente desarrollado como actividades con ruido las que estén pero debajo del valor de frecuencia. Se utiliza el registro de eventos de la prueba anterior.

3.1.2.1. Comparación de modelos

El objetivo es que las actividades de baja frecuencia y alta correlación que son agrupadas en un *cluster* por el *plugin Fuzzy Miner* de ProM como las de baja frecuencia y baja correlación que se eliminan del modelo, sean destacadas en color rojo con la personalización de *Fuzzy Miner* desde el HIS y así probar la correcta interpretación de ruido de la personalización. El modelo obtenido con ProM se presenta en la **Figura 24**.

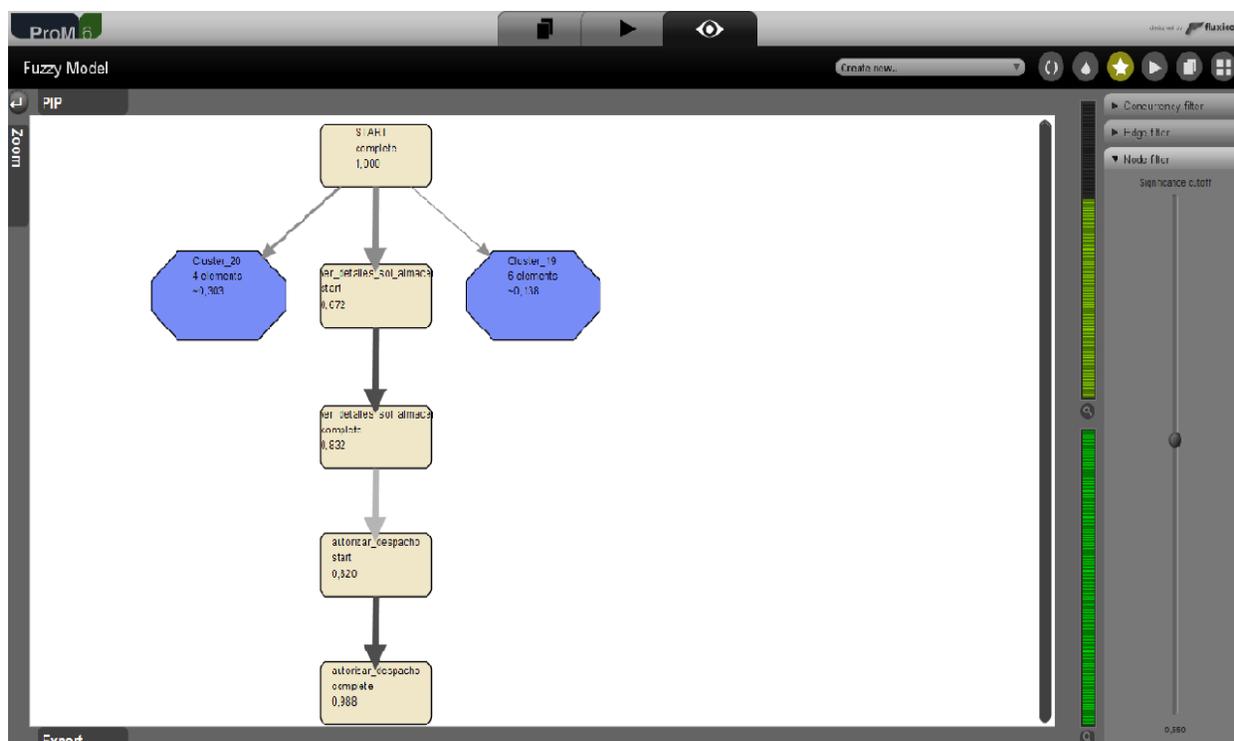


Figura 24: Interfaz de la herramienta ProM del proceso “Solicitar productos”.

Fuente: Elaboración propia.

En la **Figura 25** se muestra el modelo generado por la personalización del *plugin Fuzzy Miner* desde el HIS para el mismo valor de frecuencia (55 por ciento) con que fue generado el modelo desde el ProM y se destacan en color rojo las 12 actividades que se agrupan en los *clusters*, en este caso no se descartan actividades por lo que se puede afirmar que la personalización propuesta identificó correctamente el ruido.

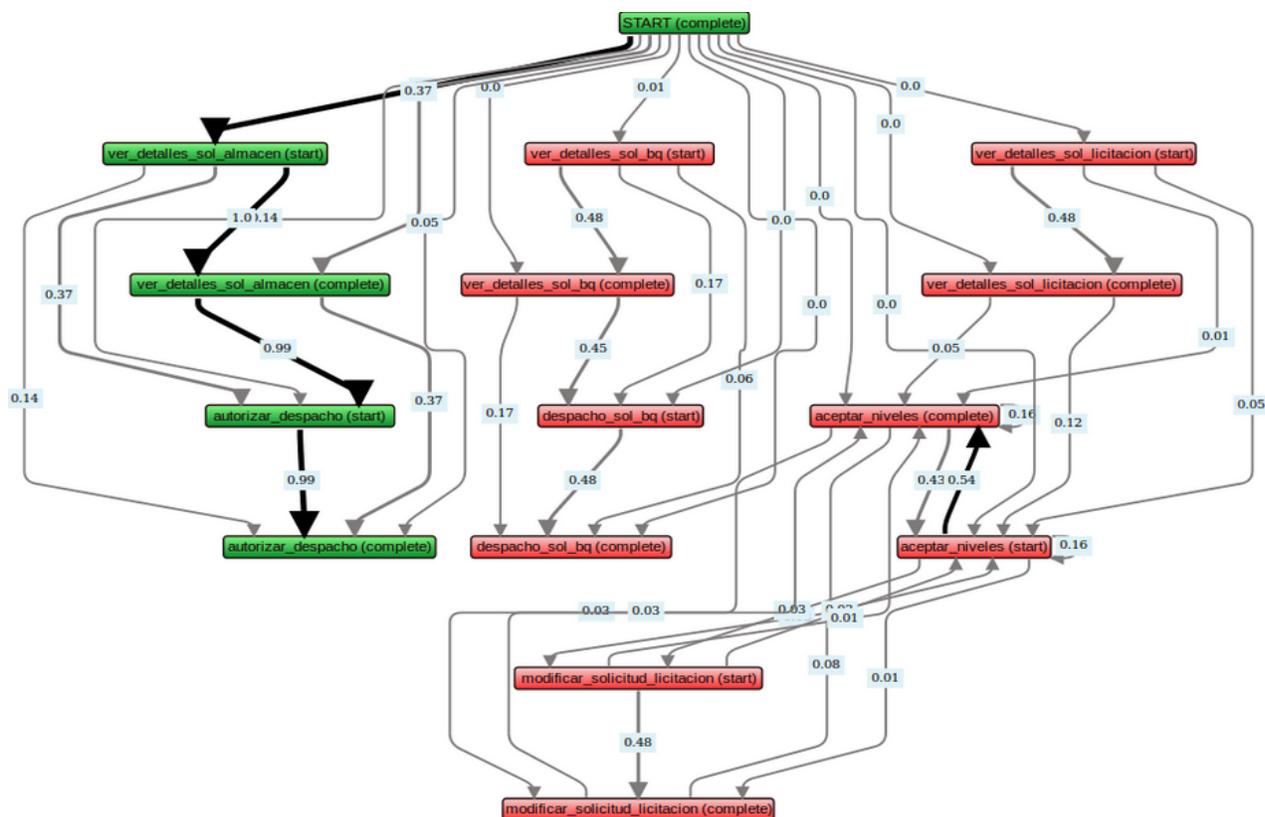


Figura 25: Interfaz del proceso “Solicitar productos” con el plugin Fuzzy Miner desde el HIS.

Fuente: Elaboración propia.

3.1.2.2. Actividades de baja frecuencia con Fuzzy Miner desde el HIS

En la **Tabla 2** se presenta como resumen de las pruebas realizadas, un *log* de cada uno de los procesos para generarlo se estableció un valor de frecuencia y conociendo con anterioridad en un análisis manual las actividades con valor de frecuencia inferior al establecido se comprobó que las mismas son visualizadas en el modelo con el color rojo.

Tabla 2: Interpretación de la presencia de ruido con datos obtenidos por el plugin Fuzzy Miner desde el HIS.

Procesos del HIS	Fechas(registro de eventos)	Valor de frecuencia(VF)	Análisis manual (actividades con valor inferior al VF)	HIS (actividades con valor inferior al VF)
Distribuir producto	2011-05-01 /2015-05-01	80%	Autorizar_despacho(start):0.73	Autorizar_despacho(start)
Desincorporar productos	2011-03-01 /2015-05-01	42%	Acta_de_desincorporacion(complete):0.35 Aprobar_desincorporacion(st	Acta_de_desincorporacion(complete) Aprobar_desincorporacion(st

			art): 0.27	on(start)
Procesar solicitudes	2012-01-01 /2015-04-01	85%	crear_contrato(start):0.23 crear_contrato(complete):0.48 crear_cronograma_de_entrega(start):0.71 crear_cronograma_de_entrega(complete):0.68	crear_contrato(start) crear_contrato(complete)) crear_cronograma_de_entrega(start) crear_cronograma_de_entrega(complete)
Transferencia hospitalaria	2011-10-05 /2015-11-01	75%	Procesar_transferencia_hospitalaria(start): 0.71	Procesar_transferencia_hospitalaria(start)
Interconsultas	2008-01-03 /2015-10-01	45%	0	0
Solicitud interconsulta hospitalaria	2011-02-01 /2015-05-05	70%	Crear_hoja_interconsulta_hospitalaria(start): 0.48 Crear_hoja_preeanestesiica(start):0.53	Crear_hoja_interconsulta_hospitalaria(start) Crear_hoja_preeanestesiica(start)

3.1.2.2.1. Ejemplo de actividades de baja frecuencia en el proceso “Solicitar productos” con Fuzzy Miner desde el HIS

Para medir la importancia de las actividades se establece como valor del filtro “Frecuencia de actividades” 30 por ciento (**Figura 26**), en un análisis manual se detectan como actividades con frecuencia por debajo a este valor el evento aceptar_niveles (complete):0.06, aceptar_niveles (start):0.06, ver_detalles_sol_bq (start):0.27; ver_detalles_sol_licitacion (start):0.26, modificar_solicitud_licitacion (start):0.17, modificar_solicitud_licitacion (complete):0.17 estas deben ser las actividades que se visualicen en color rojo en el modelo.

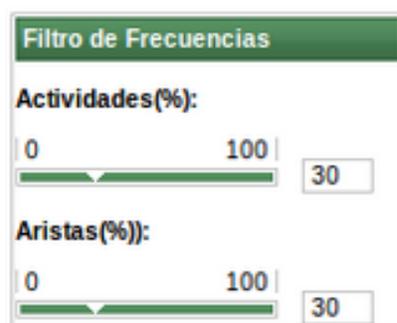


Figura 26: Interfaz del HIS para la configuración del valor Frecuencia de actividades.

Fuente: Elaboración propia.

Con el valor de frecuencia establecido en el modelo que se obtiene (**Figura 27**) se muestran como actividades de baja frecuencia, las señaladas anteriormente, por tanto se concluye que la técnica cumple con su objetivo, destacar en el modelo las actividades que el usuario considera como ruido.

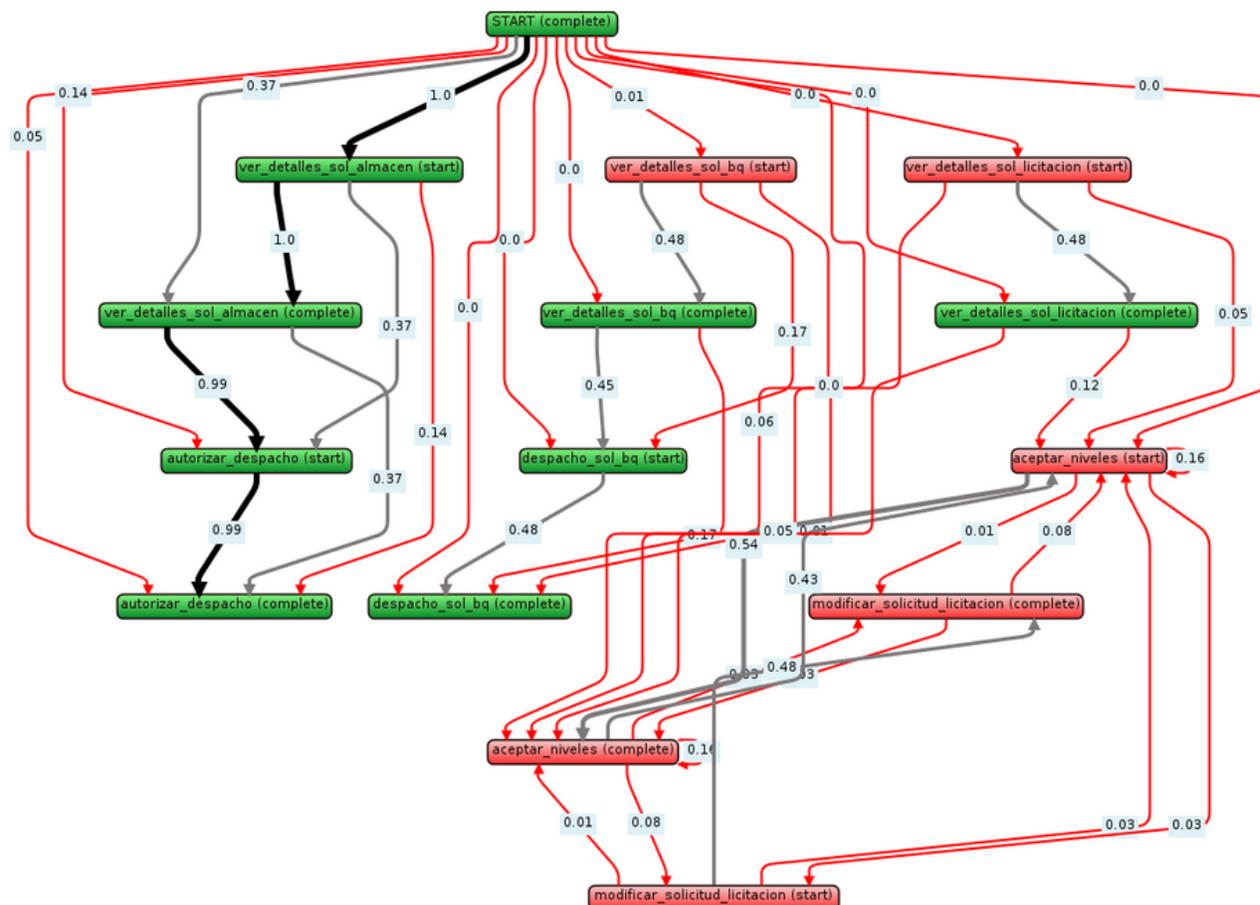


Figura 27: Interfaz de actividades con frecuencia menor a 30 por ciento.

Fuente: Elaboración propia.

3.1.3. Identificación de ruido en aristas

Para la prueba se utilizó el proceso “Desincorporar producto” del módulo almacén del HIS que cuenta con cinco actividades diferentes con un total de 16 ocurrencias. El registro de eventos recopila seis ejecuciones del proceso en las fechas comprendida entre 1 de mayo del 2011 y 12 de abril del 2015.

Como ejemplo se establece como valor del Filtro de artistas 30 por ciento con el objetivo de validar la interpretación de ruido en las aristas, se realizó un análisis manual de las relaciones entre actividades y luego se comparó con los resultados del modelo generado con la

personalización del *plugin Fuzzy Miner* como se muestra en la **Tabla 3**, arrojando como resultado que las relaciones entre las actividades detectadas con valor inferior al del Filtro de aristas son destacadas en el modelo resultante en color rojo **Figura 28**.

Tabla 3: Ejemplo de interpretación de ruido en aristas del proceso “Desincorporar Producto” del HIS.

Actividades	(1)	(2)	(3)	(4)	(5)	Total (ruido)	HIS
(1)	---	0.23	0.62	---	---	1	1
(2)	---	0.17	0.47	0.62	0.23	2	2
(3)	---	1.0	0.17	0.23	0.07	3	3
(4)	---	---	---	---	---	0	0
(5)	---	---	---	---	---	0	0

Leyenda de actividades

- (1): START (complete).
- (2): acta_de_desincorporacion (complete).
- (3): acta_de_desincorporacion (start).
- (4): aprobar_desincorporacion (start).
- (5): aprobar_desincorporacion (complete).

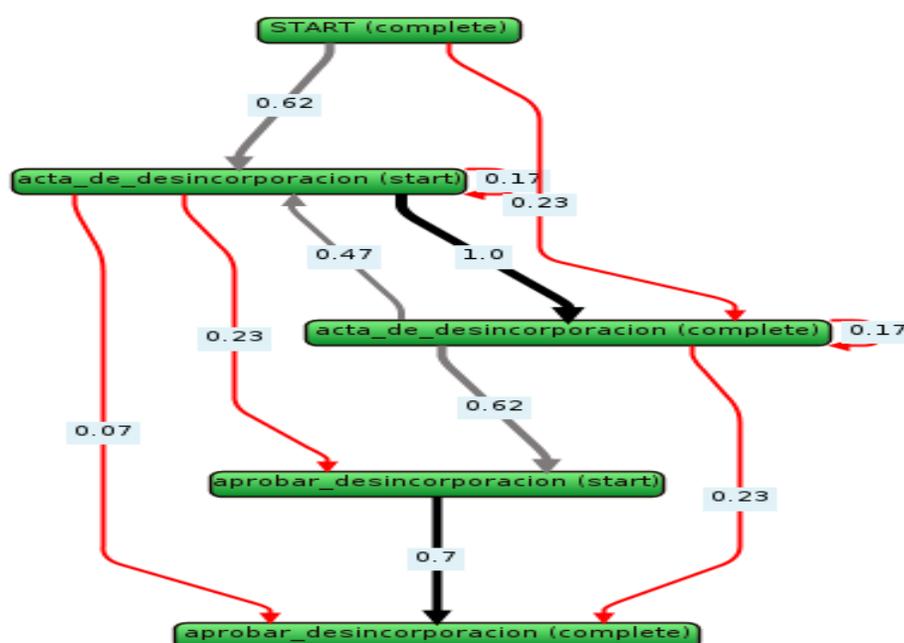


Figura 28: Interfaz del HIS de aristas con ruido en el proceso “Desincorporar producto”.

Fuente: Elaboración propia.

3.2. Conclusiones parciales

- ✓ La validación relacionada con el comportamiento de métricas del *plugin Fuzzy Miner* de ProM demuestra la necesidad de tener en cuenta todas las métricas a la hora de desarrollar una personalización para el Sistema de Información Hospitalaria si se quieren generar modelos completos y con todos los aspectos de interés.
- ✓ Las pruebas realizadas demostraron que existe una correcta interpretación del ruido en los modelos generados por la personalización lo cual resalta la efectividad de la misma en la detección del ruido en los registros de eventos.

CONCLUSIONES

El análisis de los principales conceptos, definiciones y herramientas asociadas al dominio del problema permitió:

- Elaborar el marco teórico metodológico referente a la detección de ruido en los procesos.
- Realizar un análisis comparativo del estado del arte para definir las bondades del Fuzzy Miner como propuesta de solución.

El análisis de la técnica Fuzzy Miner permitió:

- Establecer las ventajas que tiene para el tratamiento y manejo del ruido en los procesos del sistema.
- Identificar las métricas necesarias que impactan en la comprensión del modelo del proceso.

La personalización del *plugin Fuzzy Miner* de ProM en el Sistema de Información Hospitalaria del CESIM permitió:

- Detectar las actividades de baja frecuencia (ruido) en el registro de eventos según los parámetros de configuración definidos por el usuario.
- Utilizar la técnica desde el sistema, durante la etapa de diagnóstico de procesos.

La aplicación de la personalización del *Fuzzy Miner* sobre un registro de eventos del HIS permitió:

- Realizar análisis de los modelos obtenidos.
- Apoyar investigaciones asociadas a la ejecución de procesos hospitalarios.
- Utilizar la información almacenada en el sistema para generar conocimientos.

La validación realizada a partir de las pruebas definidas para la personalización desarrollada, evidencian que la propuesta de solución identifica correctamente en un registro de eventos el comportamiento considerado como ruido.

RECOMENDACIONES

Crear una funcionalidad para animar el modelo a partir del registro de evento y la mejor instancia del modelo con el objetivo de ver cómo se comportan las instancias del proceso.

Desarrollar un componente que permita la interacción de los modelos generados para contribuir a su comprensión.

REFERENCIAS BIBLIOGRÁFICAS

ÁLVAREZ DE ZAYAS, C.M. 1999. *LA INVESTIGACION CIENTIFICA EN LA SOCIEDAD DEL CONOCIMIENTO* [en línea]. S.l.: s.n. Disponible en: http://www.google.com/cu/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&uact=8&ved=0CCcQFjAC&url=http%3A%2F%2Fecaths1.s3.amazonaws.com%2Ftesis%2F48107023.UI_LIBRO%2520BASE.pdf&ei=fdlxVYCMBYmpyATd54KQBg&usq=AFQjCNFmiWzPfqUgqlaLAbNYqShM1r_nbg&sig2=UxtDZYsYBvU3OMde0QYHPA&bvm=bv.95039771,d.aWw.

BRATOSIN, C.C., SIDOROVA, N. y VAN DER AALST, W.M.P. 2010. *Distributed genetic miner*. IEEE CEC 2010. Barcelona.

BURBECK, S. 2012. Applications Programming in Smalltalk-80(TM), How to use Model-View-Controller (MVC). [en línea]. Disponible en: <http://st-www.cs.uiuc.edu/users/smarch/st-docs/mvc.html>.

CHAMORRO, M.C. y MATURANA, S., 2013. *Método para Aplicar Minería de Procesos a la Distribución de Bebestibles No Alcohólicos*. 2013.

CHRISIS, M.B., KONRAD, M. y SHRUM, S. 2009. *Guía para la integración de procesos y la mejora de productos*. segunda edición.

DE WEERDT, J. et al 2012. A multi-dimensional quality assessment of state-of-the-art process discovery algorithms using real life event logs. *Information Systems.*, vol. 37, pp. 654-676.

DEACON, J. 2012. Model-View-Controller (MVC) Architecture.

DUSTDAR, S., HOFFMANN, T. y VAN DER AALST, W.M.P. 2005. *Mining of ad-hoc business processes with TeamLog*. *Data & Knowledge Engineering*. S.l.: s.n. 55.

ECURED 2015. Patrones de diseño y arquitectura. [en línea]. [Consulta: 2 mayo 2015]. Disponible en: http://www.ecured.cu/index.php/Patrones_de_dise%C3%B1o_y_arquitectura.

FRANKY, C. 2010. Java EE 5 (sucesor de J2EE). [en línea]. Disponible en: http://www.acis.org.co/fileadmin/Conferencias/ConfConsueloFranky_Abr19.pdf.

GÜNTHER, C.W. 2009. *Process Mining in Flexible Environments*. Tesis Doctoral. The Netherlands: Eindhoven University of Technology.

ISO 9001, 2008. *NORMA INTERNACIONAL ISO 9001* [en línea]. 2008. S.l.: s.n. Disponible en: <http://farmacia.unmsm.edu.pe/noticias/2012/documentos/ISO-9001.pdf>.

JGrapht y JGraph. [en línea] 2015. [Consulta: 20 marzo 2015]. Disponible en: <http://ordenador.wingwit.com/Programacion/java-programming/89218.html>.

LAUDON, K.C. y LAUDON, J.P. 2012. *Sistemas de información gerencial*. [en línea]. 12. S.l.: s.n. [Consulta: 30 noviembre 2014]. Disponible en: <http://es.slideshare.net/jes4791/sistemas-de-informacion-gerencial-12-edicin-kenneth-c-laud-on-jane-p-laudon>.

LI, J. 2010. Fuzzy Miner Plug-in 1.2.

LIU, X. 2011. Developing applications with Facelets, JSF and JS. [en línea]. Disponible en: <http://today.java.net/pub/a/today/2006/08/29/developing-with-facelets-jsf-jsp.html>.

LYBESHARI, E., 2012. *Process mining in Intensive Care Unit Data*. 2012.

MANS, R., SCHONENBERG, M., SONG, M. y VAN DER AALST, W.M.P. 2009. Application of Process Mining in Healthcare. A Case Study in a Dutch Hospital. *Eindhoven University of Technology, Netherlands*, pp. 425-438.

MEDEIROS, A.K., WEIJTERS, A.J.M.M. y VAN DER AALST, W.M.P. 2005. *Genetic Process Mining: A Basic Approach and Its Challenges*.

MONSON-HAEFEL, R. 2004. *Enterprise JavaBeans*. 4ta. S.l.: s.n. ISBN 0-596-00530-X.

ORELLANA, A. y SÁNCHEZ, Y. 2014. Minería de Procesos en salud. Caso de Estudio: modelado de los procesos del área de Emergencia. *Twelfth LACCEI Latin American and Caribbean Conference for Engineering and Technology*.

PÉREZ, D. 2014. *Técnica para el diagnóstico de variantes de procesos de negocio*. Tesis de Maestría. Cuba: Universidad de las Ciencias Informáticas.

pgAdmin. [en línea] 2015. [Consulta: 30 marzo 2015]. Disponible en: <http://www.pgadmin.org/>.

PRESSMAN, R.S. 2002. *Ingeniería de Software, un enfoque práctico*. Quinta edición. S.l.: s.n. ISBN 8448132149.

PROCESS MINING GROUP 2009. fuzzy Minner-como usar el plugin en ProM. [en línea]. [Consulta: 24 febrero 2015]. Disponible en: <http://www.processmining.org/online/fuzzyminer>.

QPR SOFTWARE OYJ, 2011. *Automated Business Process Discovery Software QPR ProcessAnalyzer*. [en línea]. 2011. S.l.: s.n. Disponible en: <http://www.qpr.com/products/qpr-processanalyzer.htm>.

REDHAT 2010. Jboss Enterprise Application Platform 4.2. *RedHat* [en línea]. Disponible en: https://access.redhat.com/documentation/en-US/JBoss_Enterprise_Application_Platform/4.2/html/Getting_Started_Guide/index.html.

REIJERS, A., VAN DER AALST, W.M. y WEIJTERS, A.J.M.M. 2007. *Business Process Mining: An Industrial Application. Information Systems*.

RODRÍGUEZ, T., 2014. *Metodología de desarrollo para la Actividad productiva de la UCI*. 2014.

ROZINAT, A., MANS, R., SONG, M. y VAN DER AALST, W.M.. 2009. *Discovering Simulation Models. Information Systems*.

ROZINAT, A. y VAN DER AALST, W.M.P. 2008. *Conformance checking of processes based on monitoring real behavior. Information Systems*.

RUBIO, M.A., TORRES, J.M., BLANCO, J.M. y CAÑIZARES, R. 2010. Modelado y análisis de un proceso quirúrgico mediante técnicas de minería de procesos.

SALAS, A.V. y ROZINAT, Anne 2013. Minería de Procesos. *Revista de la Asociación de Técnicos de Informática*, vol. 223, pp. 1-55.

SAMPIERI, R.H. 1997. *Metodología de la investigación*. [en línea]. S.l.: s.n. [Consulta: 4 febrero 2015]. ISBN 968-422-931-3. Disponible en: http://www.academia.edu/6399195/Metodologia_de_la_investigacion_5ta_Edicion_Sampieri.

SOFTWARE AG 2011. ARIS Process Performance Manager. [en línea]. Disponible en: http://www.softwareag.com/corporate/products/aris_platform/aris_controlling/aris_process_performance/overview/default.asp.

VAN BART, D., 2012. *Process Mining in Healthcare Systems: An Evaluation and Refinement of a Methodology*. 2012.

- VAN DER AALST, W.M.P. 2011a. *Process Mining. Discovery, Conformance and Enhancement of Business Processes*. Springer Heidelberg Dordrecht. London New York: Springer. ISBN 978-3-642-19345-3.
- VAN DER AALST, W.M.P. 2011b. *Process Mining. Discovery, Conformance and Enhancement of Business Processes*. Springer Heidelberg Dordrecht. London New York: Springer. ISBN 978-3-642-19345-3.
- VAN DER AALST, W.M.P. et al. 2007. *Business process mining: An industrial application. Information Systems*.
- VAN DER AALST, W.M.P. et al. 2009. ProM: The Process Mining Toolkit. *Proceedings of BPM(Demos)* [en línea], vol. 489. Disponible en: CEUR-WS.org.
- VAN DER AALST, W.M.P. et al. 2011. *Process Mining Manifesto*. [en línea]. S.I.: IEEE. Disponible en: <https://air.unimi.it/retrieve/handle/2434/174874/182838/PMM.pdf>.
- VAN DER AALST, W.M.P. y GÜNTHER, C.W. 2007. Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics. *Business Process Management*. Alonso,G., Dadam, P. y Rosemann, M. Berlín, Alemania: Springer- Berlin Heidelberg, pp. 328–343. ISBN 978-3-540-75182-3.
- VAN DER AALST, W.M.P., MEDEIROS, A. y WEIJTERS, A.J.M.M. 2007. *Genetic Process Mining: An Experimental Evaluation. Data Mining and Knowledge Discovery*. S.I.: s.n. 14.
- VAN DER AALST, W.M.P. y WEIJTERS, A.J.M.M. 2004. Process mining: a research agenda. *ELSEVIER*, pp. 231-244.
- VAN DOREMAN, B. 2012. *Process Mining in Healthcare Systems: An Evaluation and Refinement of a Methodoregistry*. Netherlands: Eindhoven University of Technoregistry.
- VERBEEK, H.M.W., 2010. *ProM 6 Tutorial* [en línea]. 2010. S.I.: s.n. Disponible en: <http://www.promtool.org/prom6>.
- WEBSTER, C., 2011. *Conformance Checking of Processes Based on Monitoring Real Behavior..* 2011. S.I.: Trabajo presentado en EHR Workflow Inc.

WEIJTERS, A.J.M.M. y RIBEIRO, J.T.S. 2011. Flexible Heuristics Miner (FHM). *En Computational Intelligence and Data Mining (CIDM) IEEE Symposium on. IEEE*, pp. 310-317.

WESLEY, A. y STEVENS, Rob Pooley 2002. *Utilización de UML en Ingeniería del Software con Objetos y Componentes*.

YZQUIERDO, R. 2012. *Modelo para la estimación de información ausente en las trazas usadas en la minería de proceso*. Tesis de Doctorado. La Habana, Cuba: Universidad de las Ciencias Informáticas.

ANEXOS

Anexo 1. Entrevista.

- 1) ¿Cuáles son los desafíos de hoy al considerar un proyecto de Minería de Procesos en una organización?
- 2) ¿Qué características deben poseer los sistemas de información para aplicar técnicas de Minería de Procesos?
- 3) ¿Cuáles son las principales empresas que ofrecen la funcionalidad de realizar Minería de Procesos?
- 4) Se tiene entendido que la definición de ruido en el área de Minería de Procesos no es un comportamiento erróneo ¿Qué se entiende cómo ruido?
- 5) ¿Cuáles son las técnicas de descubrimiento de procesos que tratan el ruido en un registro de eventos?
- 6) Ventajas y desventajas que posee la técnica Fuzzy Miner de ProM.
- 7) ¿Qué métricas usa la técnica Fuzzy Miner para tratar el ruido?
- 8) ¿Característica de la técnica?
- 9) ¿Qué impacto trae consigo detectar y manipular la presencia de ruido en los modelos generados a partir del uso de la técnica Fuzzy Miner de ProM?
- 10) ¿Cómo mejorar la usabilidad y entendimiento de los modelos de procesos para los no expertos en Minería de Procesos?

La entrevista favoreció la comprensión de temas debido al contacto directo con conocedores de la materia, capaces de enriquecer, corroborar y refutar algunos argumentos planteados durante el análisis documental que no eran correctos, las respuestas correspondientes a cada una de las preguntas realizadas se pueden encontrar en el desarrollo de la investigación.

Perfil de los entrevistados:

Los entrevistados fueron seleccionados por ser considerados conocedores de la disciplina de investigación Minería de Procesos, tanto por su vinculación profesional, como académica, de manera que pudieron proporcionar información relevante acerca del tema.

A través de este método de investigación se han alcanzado los siguientes objetivos:

-  Obtener una visión global cualitativa del ámbito de estudio a través de las percepciones y apreciaciones de conocedores del área de Minería de Procesos.
-  Recopilar información cualitativa acerca de la emergente disciplina de investigación en cuestión.

- ✚ Posibilitar el análisis del temático objeto de estudio desde diferentes perspectivas que enriquezcan el contenido.
- ✚ Contrastar la información disponible, gracias a la diversidad de opiniones obtenidas mediante la realización de esta técnica.

Anexo 2. Requisito funcional “Generar modelo de proceso” del componente para la detección de variabilidad en procesos hospitalarios.

Tabla 4: Requisito funcional del componente para la detección de variabilidad en procesos hospitalarios.

Nº	Nombre	Descripción	Prioridad para el cliente	Complejidad	Referencias cruzadas
1.	Generar modelo de procesos	Permite obtener un modelo de procesos a partir de configurar los parámetros requeridos y seleccionar el tipo de análisis.	Alta	Alta	Ver: CESIM_SIGEC_Esp especificacion_de_casos _de_uso_Generar modelo de procesos. CU 1

Anexo 3. Especificación del caso de uso “Generar modelo de proceso” del componente para la detección de variabilidad en procesos hospitalarios.



Figura 29: Diagrama de casos de uso Generar modelo de procesos.

Fuente: Elaboración propia.

Especificación de casos de uso

CU 1. Generar modelo de procesos

Objetivo	Permite generar un modelo de proceso.
Actores	Analista de procesos hospitalarios.
Resumen	El caso de uso inicia cuando el actor accede a la opción Generar modelo de proceso, el sistema brinda la posibilidad de introducir los criterios para generar el modelo, el actor introduce los datos que considera como criterios para generar el modelo, el sistema a partir de los criterios seleccionados y consultando las entidades correspondientes genera el modelo de proceso. El caso de uso finaliza

	cuando queda generado el modelo de proceso.	
Complejidad	Alta	
Prioridad	Alta	
Precondiciones	Que existan datos de los eventos almacenados en el sistema.	
Postcondiciones	Queda generado el modelo de proceso.	
Flujo de eventos		
Flujo básico <Nombre del flujo básico>		
	Actor	Sistema
	El caso de uso inicia cuando se accede a la opción Generar modelo de procesos	
		<p>Brinda la posibilidad de introducir los criterios para generar el modelo:</p> <ul style="list-style-type: none"> • Procesos • Tipo de análisis • Desde • Hasta <p>y permite:</p> <ul style="list-style-type: none"> • Generar modelo de proceso <p>Cancelar operación. Ver Evento 1: "Cancelar operación."</p>
	Introduce o selecciona los datos que considera como criterios para generar el modelo y selecciona la opción de Generar modelo	
		<ul style="list-style-type: none"> • Si el usuario deja valores obligatorios vacíos, ver Evento 2: "Existen campos obligatorios vacíos". • Si al buscar información para generar el modelo el sistema no encuentra ninguna, ver Evento 3: "No se encontró información."
		<p>Genera el modelo a partir de los criterios seleccionados.</p> <p>Y permite:</p> <ul style="list-style-type: none"> • Exportar el modelo. Ver Evento 4: "Exportar modelo de proceso". • Ampliar el modelo. Ver Evento 5: "Ampliar modelo de proceso".
		Termina el caso de uso

Flujos alternos		
Evento 1. "Cancelar operación"		
	Actor	Sistema
	Selecciona la opción de Cancelar operación.	
		Regresa a la vista anterior.
		Termina el caso de uso.
Evento 2. "Existen datos incompletos"		
		Muestra un indicador sobre los campos incompletos.
		Regresa al paso 2 del Flujo Normal de Eventos .
Evento 3. "No se encontró información"		
1.		Muestra el mensaje de información "No se encontró información que cumpla con los parámetros seleccionados."
2.		Regresa al paso 2 del Flujo Normal de Eventos .
Evento 4. "Exportar modelo de proceso"		
1.	Selecciona la opción de Exportar	
2.		Exporta los modelos
Evento 5. "Ampliar modelo de proceso"		
	Selecciona la opción de Ampliar	
		Amplia el modelo en otra pestaña
1.		Termina el caso de uso
Relaciones	CU incluidos	No existen
	CU extendidos	

Anexo 4. Especificación del diseño de caso de prueba de "Generar modelo de proceso" del componente para la detección de variabilidad en procesos hospitalarios.

Tabla 5: Escenario Generar modelo de procesos.

Escenario	Descripción	Seleccionar procesos	Seleccionar tipo de análisis	Desde	Hasta	Respuesta del sistema	Flujo central
EC 1.1 Generar modelo de proceso	Se desea generar un modelo de procesos para analizar el mismo.	V	V	V	V	Muestra un modelo de proceso.	Muestra la interfaz para generar los modelos de procesos. Se introducen los datos. Se selecciona la opción Generar.
		<i>Solicitar Productos</i>	Perspectiva Temporal	10/11/2011	12/03/2015		
		V	V	V	V		
		<i>Solicitar Productos</i>	Incongruencias	08/10/2010	07/09/2012		
		V	V	V	V		
		<i>Procesar Solicitudes</i>	Ruido	12/05/2009	10/11/2011		
EC 1.2 Cancelar operación	Cancelar la opción de Generar modelo de proceso	NA	NA	NA	NA	Regresa a la vista anterior.	Muestra la interfaz para generar los modelos de procesos. Se introducen los datos. Se selecciona la opción Cancelar.
EC 1.2 Existen datos	Luego de haber introducido los datos, el sistema los	V	/	/	V	Muestra un indicador sobre	Muestra la interfaz para generar los
		<i>Procesar</i>	<i>Vacío</i>	<i>Vacío</i>	12/03/2015		

incompletos	verifica y valida, de haber datos incompletos, el sistema muestra un indicador sobre los campos incompletos.	<i>Solicitudes</i>				los campos incompletos.	modelos de procesos. Se introducen los datos incompletos. Se selecciona la opción Generar. Muestra un indicador sobre los campos incompletos.
		<i>I</i>	V	V	<i>I</i>		
		<i>Vacío</i>	Eventualidades	06/11/2012	Vacío		
		<i>Procesar Solicitudes</i>	Perspectiva Temporal	18/09/2015	13/02/2012		
		<i>V</i>	V	<i>I</i>	<i>I</i>		
		<i>Solicitar Productos</i>	Eventualidades	05/11/2011	02/02/2010		

Descripción de variables

Tabla 6: Descripción de las variables.

No	Nombre de campo	Clasificación	Valor Nulo	Descripción
1	Seleccionar procesos	Campo de selección	No	Se selecciona una de las opciones que muestra el sistema.
2	Seleccionar tipo de análisis	Campo de selección	No	Se selecciona una de las opciones que muestra el sistema.
3	Desde	Campo de selección	No	Se selecciona una de las opciones que muestra el sistema.
4	Hasta	Campo de selección	No	Se selecciona una de las opciones que muestra el sistema.

Anexo 5. Comportamiento de medidas y métricas.

Utilizar por defecto todas las medidas y métricas del *plugin Fuzzy Miner* de ProM en su personalización para el HIS garantiza que el modelo generado no sufra de notorias variaciones, dichas variaciones se resumen y explican a continuación a través de un ejemplo real:

Se utiliza el proceso “Solicitar productos” del módulo almacén del HIS que cuenta con 15 actividades diferentes y un total de 5713 ocurrencias. El registro de eventos recopila 1140 ejecuciones del proceso en las fechas comprendida entre 1 de mayo del 2011 y 4 de abril del 2015. Se hace uso de la herramienta ProM para mostrar el impacto que tienen las métricas al generar el modelo.

Recibiendo como entrada el registro de eventos para poder seleccionar la técnica como se muestra en la **Figura 30** y manteniendo las configuraciones por defecto de las métricas se obtiene el modelo difuso con todas las actividades y sus relaciones (**Figura 31**).

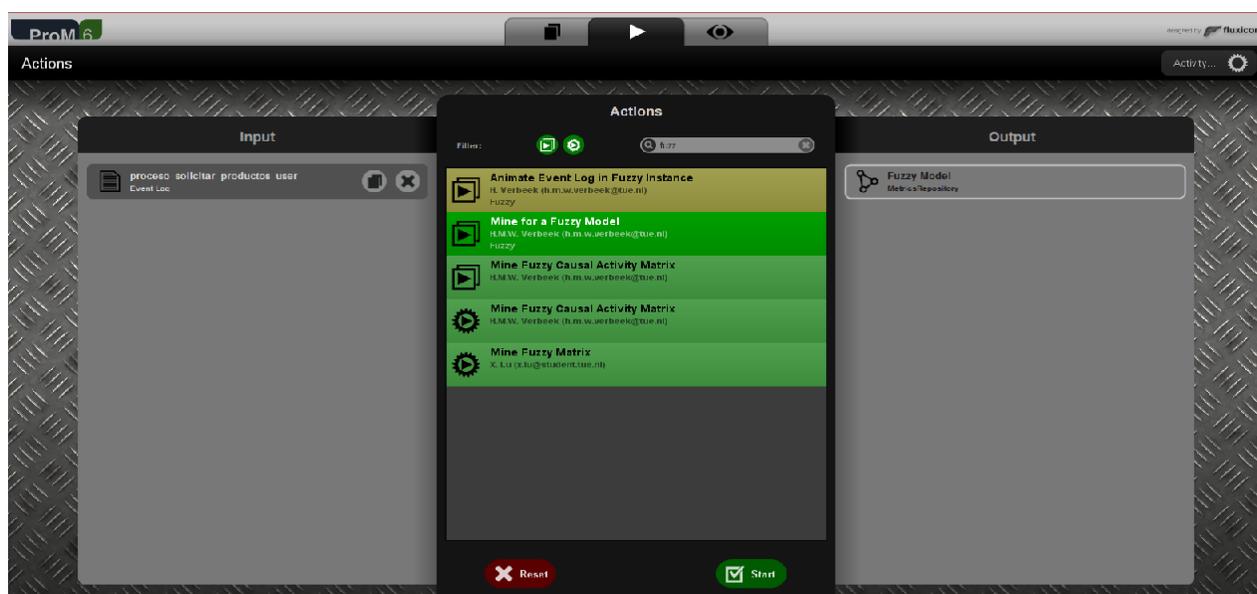


Figura 30: Interfaz de la herramienta ProM para importar registro de eventos.

Fuente: Elaboración propia.

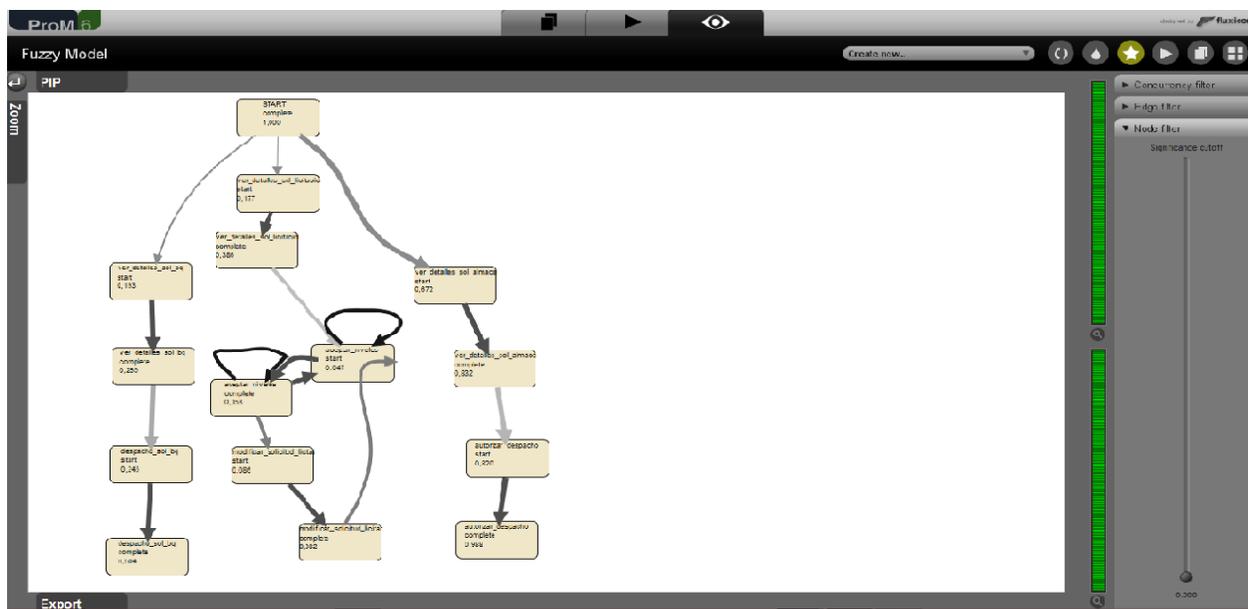


Figura 31: Interfaz de la herramienta ProM del proceso “Solicitar productos” con todas las métricas activas.

Fuente: Elaboración propia.

Si son desactivadas las métricas unarias (**Figura 32**) no se muestra el comportamiento de las actividades del registro de eventos (**Figura 33**).

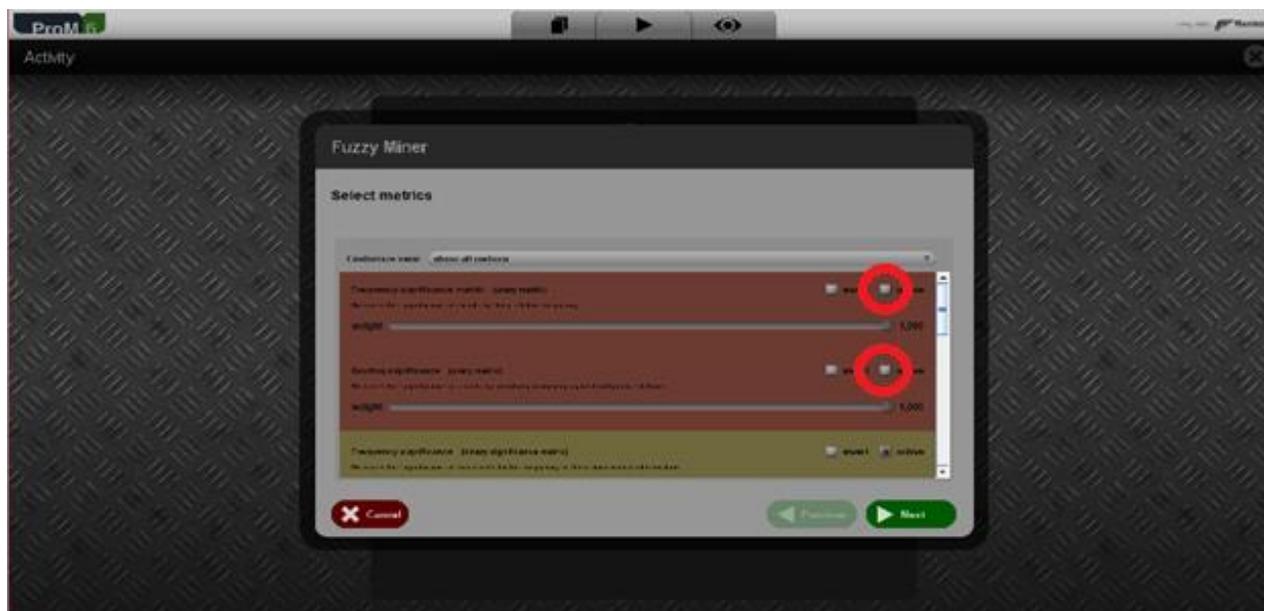


Figura 32: Interfaz de la herramienta ProM para desactivar métricas unarias.

Fuente: Elaboración propia.

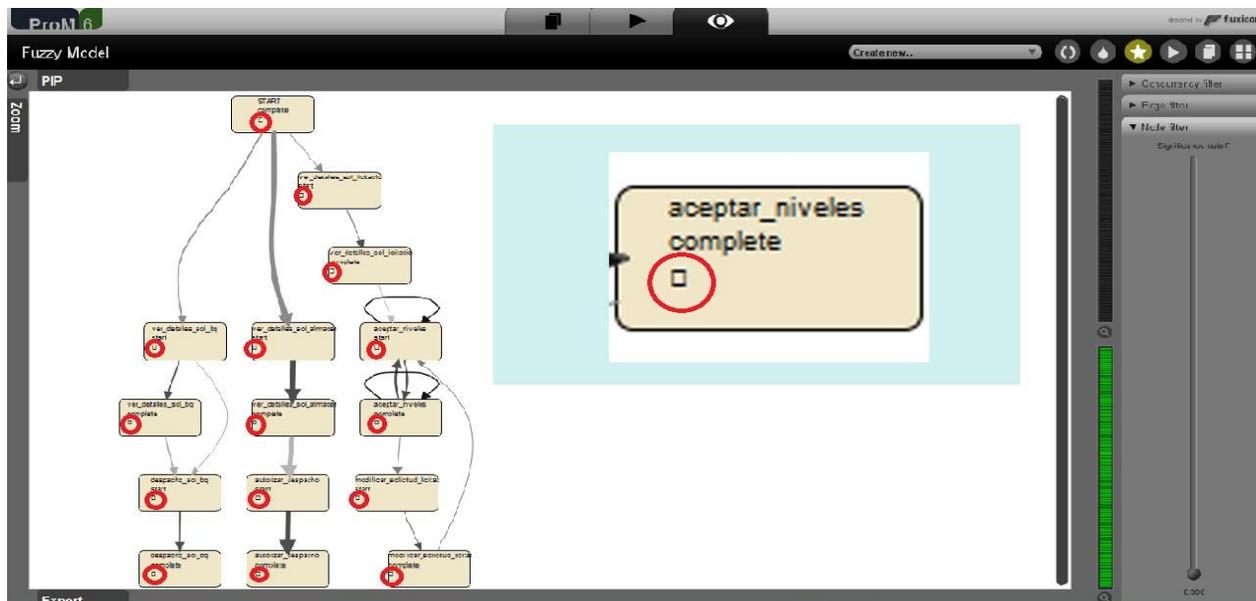


Figura 33: Interfaz de la herramienta ProM del proceso “Solicitar productos” desactivando las métricas unarias.

Fuente: Elaboración propia.

Si se desactivan las métricas binarias (**Figura 34**) se obtiene un modelo difuso (**Figura 35**) sin representar el comportamiento de las relaciones de precedencia entre actividades.

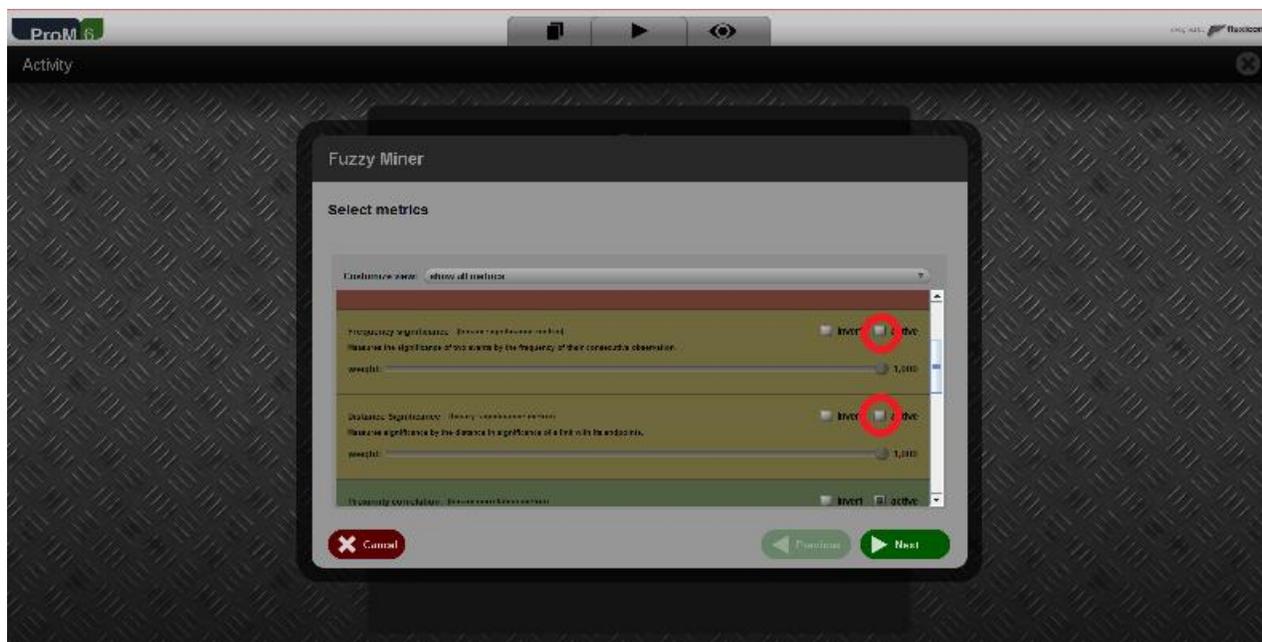


Figura 34: Interfaz de la herramienta ProM para desactivar métricas binarias.

Fuente: Elaboración propia.

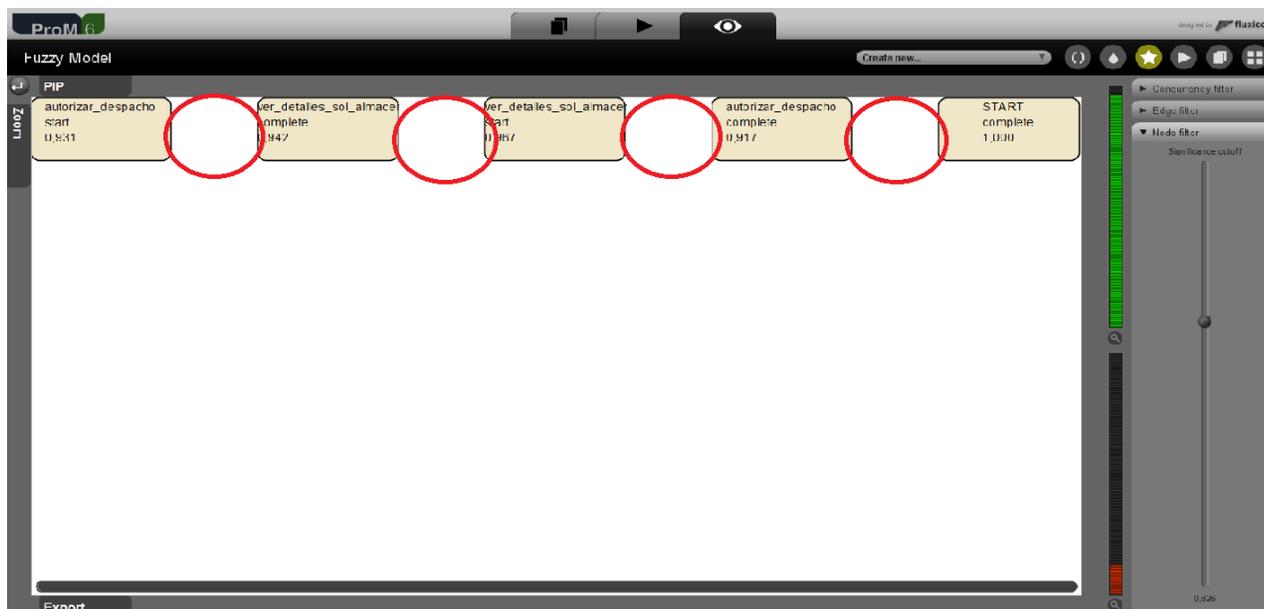


Figura 35: Interfaz de la herramienta ProM del proceso “Solicitar productos” desactivando las métricas binarias.

Fuente: Elaboración propia.

En caso de desactivar las métricas de correlación binaria (**Figura 36**) el modelo se mantiene como el de la **Figura 35**, debido a que la técnica hace tanto uso de las medidas de significación como las de correlación para generar las relaciones, de no contar con ambas, no se muestran las relaciones entre las actividades.

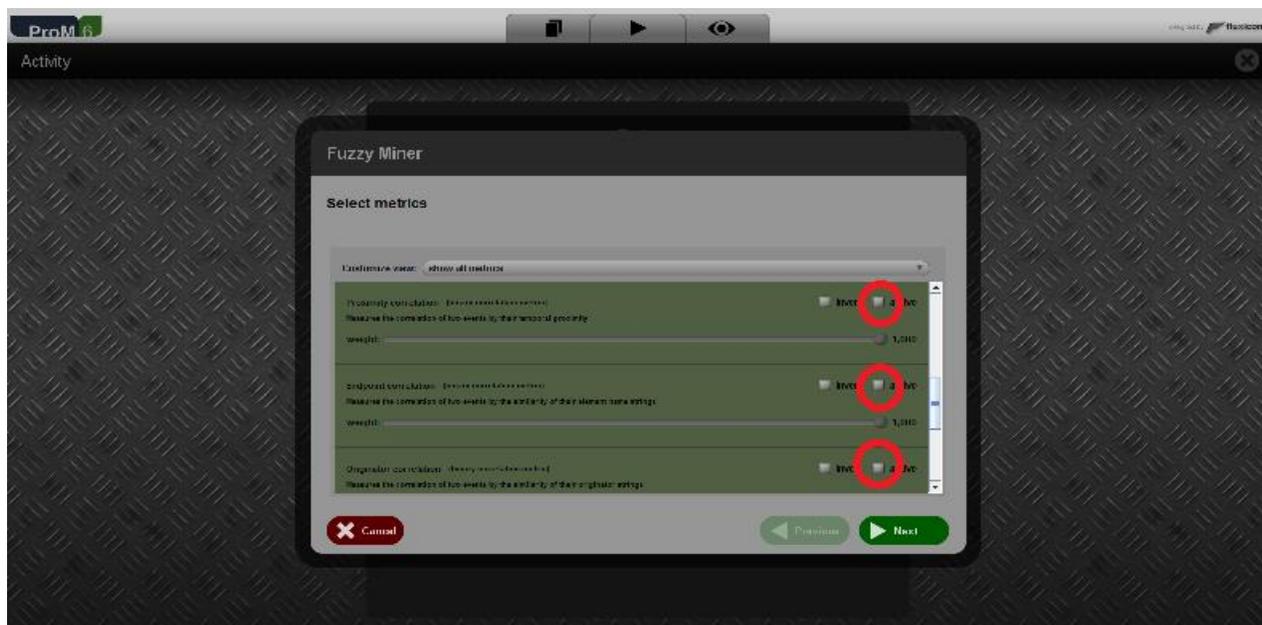


Figura 36: Interfaz de la herramienta ProM para desactivar métricas de correlación binaria.

Fuente: Elaboración propia.

Considerando los resultados obtenidos se puede afirmar que en el caso que se descarte en la personalización cualquiera de las métricas del *plugin Fuzzy Miner* de ProM, no se mostrarán todas las actividades con sus relaciones, por ende hay que tenerlas todas en cuenta para generar el modelo desde el HIS y así el usuario será quien decida las actividades y relaciones que desea visualizar.

Anexo 5. Valores para medir la frecuencia y correlación de aristas en la personalización.

Tabla 7: Valores para medir frecuencia de aristas.

FRECUENCIA	
Grosor	Valores
mayor grosor	mayores que 0.8
mediano grosor	mayores a 0.5 y menores a 0.8
poco grosor	menores que 0.5

Tabla 8: Valores para medir la correlación de aristas.

CORRELACIÓN	
Tonalidad	Valores
negro	mayores a 0.8
gris oscuro	mayores a 0.3 y menores a 0.8
gris claro	menores a 0.3