

Universidad de las Ciencias Informáticas

Facultad 7



Procedimiento de Conformidad aplicando Minería de Procesos desde la perspectiva tiempo en el Sistema de Información Hospitalaria del Centro de Informática Médica

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

Autores:

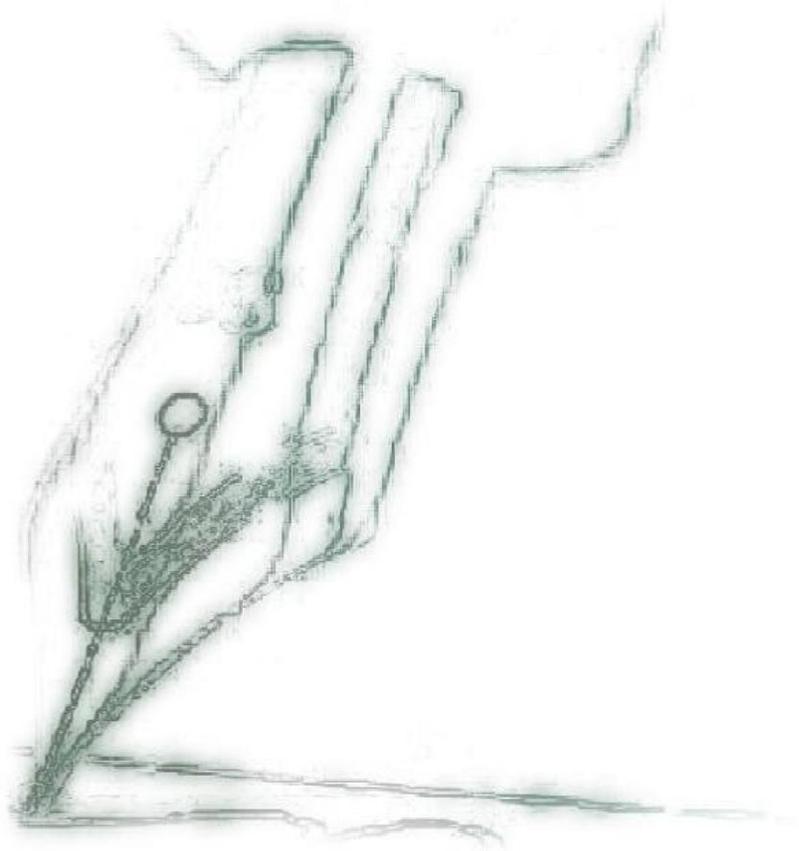
Idalberto Ricardo Cruz
Ramsés Perdomo García

Tutor: MSc. Yovannys Sánchez Corales

Co-Tutor: Ing. Arturo Orellana García

La Habana, Junio de 2014

“Año 56 de la Revolución”



"Existe una fuerza motriz más poderosa que el vapor, la electricidad y la energía atómica: la voluntad."

Albert Einstein

Idalberto

A mis padres Idalberto y Maritza, porque ellos me han acompañado en todos mis éxitos, mis fracasos, en los momentos felices y en los tristes, por sus consejos, su amor, por cuidar de nuestra familia y sobre todo por hacerme sentir orgullo de ser su hijo.

A mi hermana por ser tan linda, tan ingenua y por esos momentos en que en su amor se resume mi vida.

A mi novia Lili, tú me haces sentir especial en cada momento que paso a tu lado y eso es porque eres así especial y única.

A mis suegros por su preocupación y por acogerme como un hijo más.

A mi abuela Eva, por cuidarme y enseñarme las primeras cosas de la vida, por tantos días de dedicación y por su preocupación constante.

A mis abuelos Noldo y Lidia, por quererme en cada momento y ayudarme en los momentos difíciles.

A mami Nancy, tía Margot, tía Marlenis, tío Rolando y tío Negro, porque más que un sobrino he sido para ellos un hijo.

A mis primos por tantos años de compañía y ayuda, a ustedes más que primos los puedo llamar hermanos.

A mi mejor amigo Oscar Leandro, por poder contar con él en todo momento.

A mis compañeros de apartamento y aula, amigos de aquí y de allá, Medero, Raúl, Ardelio, Rober, Migue, Asiel, Karina y todos los que me ayudaron en algún momento.

A todos los profesores que me han ayudado a superarme.

A mis tutores Yovannys y Arturo por su paciencia, su tiempo y sus experiencias, sin las cuales hubiera sido imposible este momento.

A mi compañero de tesis, Ramsés, sin el cual tantas horas de trabajo serían inútiles, gracias mi hermano por tanto apoyo y amistad.

Ramsés

- A mi madre Eloísa, por ser mi mejor amiga y mi confidente, por ser madre, padre, tía, hermana, por ser tan brillante y ser mi orgullo, mamá me siento orgulloso de ser tu hijo.**
- A mi abuela Teresa, por cumplir todos mis caprichos, por sacrificarse tanto por mí, y sobre todo por ser la mejor abuela del mundo y entregarme su vida desde el día en que nació.**
- A mi padre Adalberto, que aunque la vida no quiso estuviese aquí, sé que me acompañará y protegerá siempre.**
- A mi abuela Águeda, por preocuparse tanto por mí y apoyarme siempre con mis estudios, gracias abuela.**
- A mis mejores amigos Hiram y Gabriel, por ser más que amigos mis hermanos, y haberme ayudado en los momentos más difíciles de mi vida, nunca lo olvidaré.**
- A mi tío Oscarito, por haberme cuidado todos estos años y haberme apoyado siempre, gracias por ser mi segundo papá.**
- A mi tío Oscar, por haberme sacado de un millón de aprietos y preocuparte siempre por mí, gracias tío.**
- A mi tía Alina, por preocuparse siempre por mi comida de la escuela y por quererme tanto y cocinarme tan rico.**
- A mi tía Niurka, por todos los cabos que me tiró para que pasase bien cómodo la escuela, gracias tía linda.**
- A mi prima Claudia por su ayuda y caerme tanto atrás en la tesis.**
- A todos mis primos, por ser mis ejemplos a seguir y apoyarme siempre en todo.**
- A mi novia por amarme tanto y aguantarme todos estos años en la universidad.**
- A Carmela, la mejor profesora que he tenido en toda mi vida, profe sin usted tal vez hoy yo no estaría aquí, gracias.**
- A mi tutor, por aguantarnos todos estos meses y ayudarnos a ser mejores.**
- A mi co-tutor por pasar varias noches ayudándonos con el documento y apoyarnos siempre en todo, gracias hermano, mucha gracias.**
- A Idalberto, mi compañero de tesis, por ser un gran amigo y compañero, te quiero mi hermano.**

Idalberto

A mi mamá, Maritza, por ser mi amiga, mi confidente, por quererme complacer en todo y sobre todo por existir.

A mi papá, Idalberto, en hacerte sentir orgulloso, va todo mi esfuerzo.

A mi hermanita, porque no hay día que pase que no quiera abrazarla.

A mi novia Lili, por estos dos hermosos años a tu lado.

A mi familia y amigos, que me han ayudado a superarme y ser una mejor persona.

Ramsés

A mi madre Eloísa.

A mis abuelas Teresa y Águeda.

A mi novia por aguantar estos años a mi lado y apoyarme a pesar de todo.

A mi familia por apoyarme y alentarme a seguir hacia adelante siempre.

A todos mis amigos que sin ellos esto no hubiese sido tan divertido.

Resumen

La minería de procesos constituye una disciplina novedosa para extraer información útil y oculta del funcionamiento de los Sistemas de Gestión de Información Modernos. Sirve como herramienta para detectar errores e inconsistencias a partir de los registros de trazas que implementan estos sistemas.

Sin embargo, en la actualidad los datos almacenados en el registro de trazas del Sistema de Información Hospitalario (HIS por sus siglas en inglés) no se aprovechan, por lo que se desconoce si el comportamiento del sistema es tal como, se modeló inicialmente. Esto permitiría identificar errores de programación que puedan comprometer el funcionamiento del sistema.

Entre los autores más destacados en el estudio de la aplicación de técnicas de conformidad en el área de la salud, se pueden citar a Bozkaya, van der Aalst y Jans. Los cuales, destacan los beneficios y desafíos al aplicar esta disciplina en un sector crítico de la sociedad, como lo es la atención sanitaria.

En la presente investigación se realizó el chequeo de conformidad al proceso Solicitar producto del módulo Almacén perteneciente al HIS. Se escogió este proceso por ser el que presenta la mayor cantidad de datos almacenados, además de que utiliza el motor de flujo de trabajo JBPM.

La validación mediante el método de fitness, permitió identificar lo cercano que están los modelos generados de la realidad. Por otro lado la aplicación de las técnicas de conformidad permitió identificar patrones de comportamiento erróneo en los datos que pueden significar fallas en el sistema. Además, se detectó la poca ejecución de varias actividades en el proceso y la poca consistencia en el comportamiento del proceso.

Palabras claves: Gestión basada en procesos, minería de proceso, registro de eventos, chequeo de conformidad, modelos, sistemas de información.

Abstract

Process mining is a novel discipline destined to extract useful information which is hidden in modern information management systems. It works as a tool to detect errors and inconsistencies in trace registries.

However, today data stored in the Hospital Information System's trace registry are not taken advantage of, and thus, the system's behavior does not match to the initial model. This could allow to identify programming errors which may compromise the system's performance.

Among the most prominent authors in the study of the application of conformance checking techniques to the area of health, it can be mentioned to Bozkaya, van der Aalst and Jans. Which highlight the benefits and challenges in applying this discipline in a critical sector of society, such as health care.

In the present investigation the conformance checking process requests belonging to the HIS Product Warehouse module was performed. This process was chosen to be the one with the greatest amount of stored data, and using the JBPM workflow engine work.

Validation by the method of fitness, fencing it possible to identify models that are generated from reality. On the other hand the application of techniques allowed us to identify patterns accordance malfunction in the data can mean failure in the system. Besides the low implementation of various activities in the process and the lack of consistency in the behavior of the process was detected.

Keywords: *Process based management, process mining, event registro, conformance checking, models, information systems.*

Tabla de Contenido

| | |
|---|----|
| INTRODUCCIÓN..... | 10 |
| CAPÍTULO 1. FUNDAMENTOS TEÓRICOS SOBRE LA MINERÍA DE PROCESOS..... | 15 |
| 1.1 Gestión basada en procesos | 15 |
| 1.2 Minería de Procesos | 16 |
| 1.2.1 Perspectivas para aplicar Minería de procesos | 18 |
| 1.2.2 Registro de eventos..... | 19 |
| 1.2.3 Clasificación de los procesos..... | 22 |
| 1.2.4 Técnicas de minería de procesos | 23 |
| 1.2.5 Métricas de calidad..... | 27 |
| 1.3 Modelos para aplicar técnicas de minería de procesos..... | 28 |
| 1.3.1 Modelo de Rozinat..... | 28 |
| 1.3.2 Modelo de Bozkaya..... | 30 |
| 1.3.3 Modelo de van der Aalst..... | 31 |
| 1.3.4 Valoración de los modelos anteriores | 33 |
| 1.4 Herramientas | 34 |
| 1.4.1 Sistemas Gestores de Bases de Datos..... | 34 |
| 1.4.2 Clientes de administración de bases de datos..... | 35 |
| 1.4.3 Herramientas de Minería de Procesos..... | 36 |
| 1.4.4 Herramientas para la extracción de registros de eventos | 36 |
| CAPÍTULO 2. PROCEDIMIENTO PARA EL CHEQUEO DE LA CONFORMIDAD APLICANDO TÉCNICAS DE MINERÍA DE PROCESOS..... | 39 |
| 2.1 Características del modelo de van der Aalst..... | 39 |
| 2.2 Descripción de las etapas y fases del modelo de van der Aalst | 40 |

| | |
|---|-----------|
| 2.2.1 Etapa 0 - Planificar y justificar. | 40 |
| 2.2.2 Etapa 1 - Extracción. | 40 |
| 2.2.3 Etapa 2 - Crear el modelo de flujo de actividades y conectar con el registro de eventos. | 47 |
| 2.2.4 Etapa 3 - Crear un modelo de proceso integrado. | 51 |
| 2.2.5 Etapa 4 - Apoyo a las operaciones. | 51 |
| CAPÍTULO 3: APLICACIÓN DEL PROCEDIMIENTO DE CONFORMIDAD AL MÓDULO ALMACEN. | 53 |
| 3.1 Módulo Almacén del Sistema de Información Hospitalaria. | 53 |
| 3.2 Aplicación del procedimiento para el chequeo de la conformidad. | 54 |
| 3.2.1 Etapa 0 - Planificar y justificar. | 54 |
| 3.2.2 Etapa 1 - Extracción. | 54 |
| 3.2.3 Etapa 2 - Crear el modelo de flujo de actividades y conectar con el registro de eventos. | 56 |
| 3.2.4 Etapa 3 - Crear un modelo de proceso integrado. | 63 |
| 3.2.5 Etapa 4 - Apoyo a las operaciones. | 66 |
| CONCLUSIONES. | 71 |
| RECOMENDACIONES. | 72 |
| REFERENCIA BIBLIOGRÁFICA. | 73 |
| GLOSARIO DE TÉRMINOS. | 77 |
| ANEXOS. | 78 |

INTRODUCCIÓN

El alto grado de desarrollo que alcanzan las Tecnologías de la Información y las Comunicaciones (TIC) (Machío, 2010), exige al ser humano utilizar su ingenio para resolver los múltiples problemas que le plantea la sociedad moderna. El proceso de informatización de la sociedad a nivel mundial, significa la aparición de múltiples retos a las aplicaciones y sistemas informáticos desplegados en las más variadas esferas de la actividad humana. El enfoque basado en procesos permite enfrentar algunos de estos retos mejorando la gestión y el desempeño dentro de la organización.

Algunas instituciones que reconocen las oportunidades que ofrecen las herramientas utilizadas en aras de mejorar los procesos de negocio (Bose & van der Aalst, 2011); implementan mecanismos para registrar la ejecución real de estos, tales como los registros de eventos, registros de trazas, bitácoras o registro (en lo adelante registros de trazas). Como alternativa de explotación de los registros de trazas surge la Minería de Procesos.

La Minería de Procesos (Rozinat & van der Aalst, 2008; van Doremalen, 2012; Bratosin, 2011) es una disciplina de investigación que se ubica entre la inteligencia computacional y la minería de datos, por una parte, y modelación y análisis de procesos, por otra (Van der Aalst, 2011). Permite extraer información oculta y útil de los registros de trazas almacenados por sistemas de información. Además, posibilita entender cómo son ejecutados en realidad los procesos en el sistema (van Doremalen, 2012). Su aplicación ayuda a identificar cuellos de botella, anticipar problemas, registrar violaciones de políticas, recomendar contramedidas, y simplificar procesos (Van der Aalst, 2011) para la mejora del funcionamiento del negocio.

Las herramientas utilizadas en la Minería de Procesos, necesitan como entrada un registro estructurado simplificado, que elimine las incongruencias existentes en los registros de trazas. Además, se debe conservar sólo la información que sea de interés para la investigación, a este nuevo registro que se estructura en una única tabla se le conoce como registro de eventos.

Existen tres tipos fundamentales de técnicas de minería de procesos: el Descubrimiento de Procesos, la Verificación o Chequeo de Conformidad y el Mejoramiento de Procesos (Van der Aalst, 2011). Las técnicas de Descubrimiento de Proceso consisten en determinar, a partir de un registro de eventos, un modelo de proceso que represente el comportamiento real del mismo. Esta técnica es la más

destacada, porque normalmente el modelo descubierto representa un modelo de procesos aunque podría describir otras perspectivas como una red social.

Las técnicas de Verificación de Conformidad necesitan un registro de eventos y un modelo como entrada. Su objetivo es mostrar las diferencias y elementos en común entre el modelo y el registro de eventos antes mencionados. Por último, las técnicas de Mejoramiento también necesitan como entrada un registro de eventos y un modelo y la salida es un modelo mejorado o extendido.

Entre las perspectivas (van der Aalst et al., 2011) que incluyen la Minería de Procesos se destacan la referente al control de flujo, a los casos, al tiempo, entre otras. La perspectiva de control de flujo se enfoca en el orden de ejecución de las actividades. El objetivo de explorar esta perspectiva es encontrar una buena caracterización de todos los caminos posibles. El resultado se expresa típicamente en términos de una red de Petri o alguna otra notación de procesos como BPMN o UML. La perspectiva de casos se enfoca en caracterizar los casos por su ruta en los procesos, los actores que trabajan en él o el valor de los datos de sus elementos. La perspectiva tiempo puede ser aplicada cuando se tiene información con relación al tiempo en el registro de eventos, permitiendo integrar al modelo las marcas de tiempo.

Las diferencias en un intervalo de tiempo entre las actividades ejecutadas por los usuarios y las relacionadas causalmente, se pueden utilizar para realizar operaciones sobre los datos (Harmen, 2012). Estas aplicaciones son especialmente efectivas y útiles en sectores críticos (van der Aalst et. al, 2011) como el de la salud (Mans et al., 2008), pues determinan el tiempo promedio de atención a un paciente, las diferencias de tiempo dentro de un proceso asistencial para un grupo de pacientes con el mismo diagnóstico y las actividades dentro del hospital pueden representar cuellos de botella, ya sea por tardanza en la atención a los pacientes o por sobrecarga en los servicios médicos.

El Sistema de Información Hospitalario (*Hospital Information System*, HIS por sus siglas en inglés) que se desarrolla en el Centro de Informática Médica (CESIM) de la Universidad de las Ciencias Informáticas (UCI), tiene como objetivo informatizar la gestión en el nivel de atención secundario. Para su funcionamiento se divide en varios módulos, entre los que se encuentra el módulo Almacén que se encarga de suministrar los productos sanitarios y los fármacos necesarios para una correcta atención al paciente, así como los materiales en condiciones óptimas de uso.

El módulo Almacén consta de varios procesos, entre los que se encuentran *Recepcionar producto*, *Despacho de producto*, *Control de stock* y *Solicitar producto*. Éste último consiste en realizar solicitudes a otro almacén de la misma institución para el reabastecimiento de determinados productos. Dicho proceso es implementado utilizando el lenguaje Java para ejecutar los procesos de negocio que son descritos mediante la Notación para el Modelado de Procesos de Negocio (BPMN por sus siglas en inglés) (Rozinat & van der Aalst, 2006), este motor de flujo de trabajo conocido como JBPM toma como entrada descripciones gráficas de los procesos.

El HIS posee un registro de trazas que almacena las actividades del proceso Solicitar producto y de los restantes procesos que se ejecutan en el sistema. Actualmente no se realiza un procedimiento de análisis y aprovechamiento de los datos almacenados en este registro, por tanto se desconoce si el comportamiento del sistema se corresponde con el proceso inicial, lo cual imposibilita encontrar diferencias y elementos en común entre el modelo y el registro de eventos.

Este registro almacena información temporal de las actividades, como fecha de inicio, fecha de fin y duración de las mismas. Entre las operaciones que se pueden realizar con esta información está calcular la duración media y el tiempo total de las actividades, así como la desviación típica. La ausencia de estas operaciones sobre la información que se encuentra en el registro de eventos vinculada al proceso Solicitar producto se imposibilita al agregar tiempo de espera estimado en la ocurrencia de las actividades, como la que transcurre desde el inicio hasta el fin de una solicitud y entre solicitudes de productos.

La inexistencia de las operaciones antes mencionadas imposibilita la detección de cuellos de botella, que puedan ralentizar el funcionamiento del sistema, ocasionados por la concurrencia de varias solicitudes de productos al mismo tiempo o por la tardanza en la realización de estas. Además de monitorizar la utilización de los recursos en el tiempo como los implementos quirúrgicos, los destinados a enfermería y los de consulta externa, para detectar funcionamientos anómalos, como la desviación de recursos, el exceso en el tiempo de estancia de un producto en el almacén o el faltante de algún producto, de forma tal que permita solicitar un producto con mayor o menos frecuencia.

Teniendo en cuenta la situación anterior, como **problema científico** se define ¿Cómo obtener un patrón de comportamiento de los procesos desde la perspectiva tiempo del módulo Almacén del Sistema de Información Hospitalaria a partir del registro de eventos del sistema?

El **objeto de estudio** es la gestión basada en minería de procesos en los sistemas de información hospitalarios y como **campo de acción** las técnicas de conformidad con respecto al tiempo en el módulo Almacén del Sistema de Información Hospitalaria.

Se definió como **objetivo general** realizar un análisis de conformidad desde la perspectiva tiempo al módulo Almacén del Sistema de Información Hospitalaria perteneciente al Centro de Informática Médica.

Las **tareas de investigación** que se definieron para darle cumplimiento al objetivo planteado fueron las siguientes:

1. Establecimiento de los fundamentos teóricos-metodológicos para la aplicación de Cheques de Conformidad a través de la Minería de Procesos.
2. Selección de las herramientas a utilizar.
3. Selección de un modelo que permita aplicar las técnicas de conformidad al módulo Almacén.
4. Aplicación de las distintas técnicas relacionadas con la conformidad que arroja la herramienta ProM a través del método diseñado.
5. Validación de la propuesta de solución, mediante el método de fitness.

Los **métodos científicos** utilizados para desarrollar la investigación (Hernández, 2011) fueron:

Métodos teóricos:

- Análisis-síntesis: para el estudio de las principales técnicas y herramientas existentes utilizadas en la aplicación de Minería de Procesos.
- Deductivo-inductivo: para analizar y definir los algoritmos a utilizar a partir de la información que brinde el registro de eventos.
- Histórico-lógico: utilizado para analizar el surgimiento, trayectoria y evolución de la Minería de Procesos para registrar una mejor comprensión del objeto y campo de estudio.
- Modelación: se utiliza para la creación y el análisis las diferentes propuestas de solución que arrojan las herramientas y tomar la que más se adecua a la investigación.

Métodos empíricos:

- Observación: como instrumento para adquirir conocimiento sobre el campo de estudio a través de la investigación directa de las herramientas.

El documento está dividido en tres capítulos, donde se explican elementos teóricos sobre Minería de procesos, así como la aplicación y validación de las técnicas de minería de procesos utilizadas en la verificación de la conformidad. Además cuenta con Introducción, Conclusiones, Recomendaciones, Referencias Bibliográficas, Glosario de Términos y Anexos.

El **Capítulo 1. Fundamentos Teórico-Methodológicos sobre la minería de procesos**, trata los conceptos fundamentales sobre la minería de procesos: perspectivas, tipos, técnicas, estado del arte y modelos; así como las herramientas a utilizar. Igualmente se analizan varios modelos de diferentes autores para aplicar las técnicas de conformidad.

El **Capítulo 2. Procedimiento para el Chequeo de la Conformidad aplicando técnicas de minería de procesos**, expone el análisis a los modelos existentes y la selección del que más se ajusta a la investigación y se describen las fases y pasos que guían la utilización de las técnicas y algoritmos de minería de procesos para el Chequeo de la Conformidad en los sistemas de información de salud.

Por último, en el **Capítulo 3. Aplicación del procedimiento de conformidad al módulo Almacén**, se valida la aplicación de las técnicas de conformidad, mediante la puesta en práctica del modelo seleccionado al módulo Almacén y se muestran los resultados obtenidos.

CAPÍTULO 1. FUNDAMENTOS TEÓRICOS SOBRE LA MINERÍA DE PROCESOS.

En el presente capítulo se abordan los conceptos fundamentales sobre la gestión basada en procesos, una breve reseña, las perspectivas, el uso en los diferentes sectores de la sociedad y los tipos de minería de procesos existentes. También un análisis de los estudios existentes tanto en Cuba como en el mundo sobre esta disciplina.

1.1 Gestión basada en procesos

Según la serie de normas internacionales ISO 9000 (2005) se define un proceso como “conjunto de actividades mutuamente relacionadas o que interactúan, las cuales transforman elementos de entrada en resultados”. Autores como (Recker et al., 2007) lo definen como un conjunto de actividades parcialmente ordenados destinadas a alcanzar un objetivo.

Mientras (Barros, 2012) define que un proceso “es un conjunto de tareas lógicamente relacionadas que existen para conseguir un resultado bien definido dentro de un negocio; por lo tanto, toman una entrada y le agregan valor para producir una salida. Los procesos tienen entonces clientes que pueden ser internos o externos, los cuales reciben a la salida, lo que puede ser un producto físico o un servicio. Estos establecen las condiciones de satisfacción o declaran que el producto o servicio es aceptable.

El enfoque basado en procesos es un concepto que aparece en los años 90 del pasado siglo, con el objetivo de hacer más eficaz el funcionamiento de las organizaciones de información. Además de aumentar el grado de satisfacción de los clientes. Con este enfoque se puede dar seguimiento a los procesos, permitiendo detectar errores y redundancias, así como gestionar los procesos interrelacionados.

La Gestión de Procesos de Negocio (*Business Process Management*, BPM por sus siglas en inglés) permite utilizando métodos, técnicas y software, diseñar, ejecutar, controlar y analizar procesos operacionales que involucran personas, organizaciones, aplicaciones, documentos y otras fuentes de información (van der Aalst, Ter Hofstede & Weske, 2003). También es conocida como una gestión integral que promueve la eficacia empresarial y la eficiencia mientras se esfuerza por innovación, flexibilidad, y la integración con la tecnología (Gao, 2013). Además está creciendo como una disciplina,

donde las nuevas tecnologías están emergiendo rápidamente, manteniendo el centro del escenario BPM en los dominios de negocio y tecnología.

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 a todos los interesados un entendimiento de la organización y su desempeño. Esto ayuda en la automatización de las actividades, la colaboración, la integración con otros sistemas, la integración de los socios a través de la cadena de valor, entre otros.

1.2 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). Esta disciplina provee un puente importante entre la minería de datos y la modelación y análisis de procesos de negocios. Su objetivo fundamental es descubrir, monitorear y mejorar los procesos reales (van der Aalst et al., 2011) a través de la extracción del conocimiento de los registros de trazas ampliamente disponibles en los actuales sistemas de información.

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 información de estos sistemas se almacenan en registros de trazas y un correcto análisis de estos datos puede ayudar a mejorar la forma en que se ejecutan y gestionan los procesos en la organización.

La minería de procesos incluye el descubrimiento, utilizado para extraer modelos de procesos a partir de un registro de eventos, la verificación de conformidad para monitorear desviaciones al comparar el modelo y el registro de eventos y la mejora de los modelos existentes (van der Aalst et al., 2011). Estos son los tres tipos de minería de procesos que se utilizan partiendo de un registro de eventos (Figura 1.1), donde cada evento se refiere a una instancia de procesos y una actividad. Los eventos pueden contener información adicional como marcas de tiempo o datos sobre recursos, lo que permite enriquecer los modelos generados.

El **descubrimiento** consiste en tomar un registro de eventos y generar un modelo sin utilizar ninguna información previa (van de Aalst, 2011). Esta técnica es la más destacada y utilizada en la actualidad, permite descubrir cómo se ejecutan realmente los procesos en la organización.

En la **verificación de conformidad** se compara un modelo de proceso existente con un registro de eventos del mismo proceso (Rozinat & van der Aalst, 2008). Esto permite comprobar si existen desviaciones entre la ejecución real y el modelo previo, utilizando estas desviaciones se realiza la **mejora** al modelo incorporándole la información obtenida en las etapas anteriores.

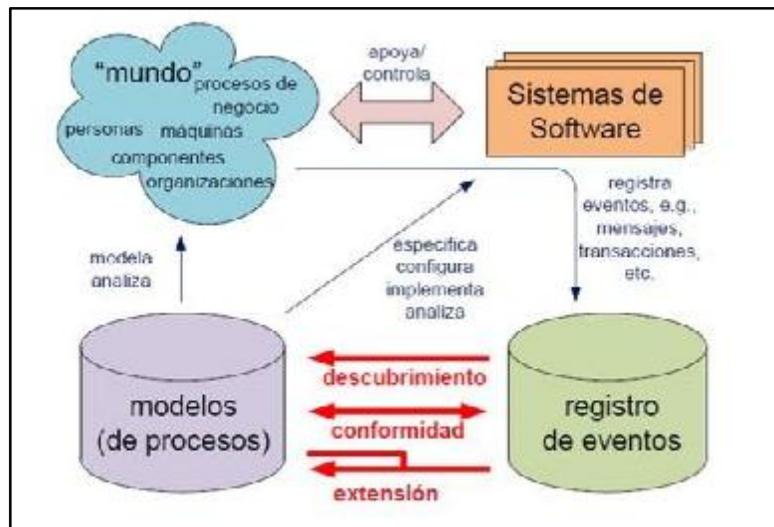


Figura 1.1. Tipos de minería de procesos

Fuente: van der Aalst, 2011

Las técnicas de minería de procesos forman parte de las tecnologías inteligentes (Seguel, 2008), por lo que se considera una herramienta dentro de la Inteligencia de Procesos de Negocios (*Business Process Intelligence*, BPI por sus siglas en inglés). El propósito de BPI es apoyar una mejor toma de decisiones empresariales (Castellanos et al., 2009), lo que se traduce en mejorar la eficiencia de los procesos de negocios y reaccionar rápidamente a los cambios del mercado con el fin de ser competitivos en la dinámica era de la Internet.

En el año 2009, se funda la Fuerza de Trabajo del Instituto de Ingenieros Eléctricos y Electrónicos (*Institute of Electrical and Electronic Engineers*, IEEE por sus siglas en inglés), con el objetivo de promover la investigación, desarrollo, educación, implementación, evolución y entendimiento acerca de la minería de procesos (van der Aalst et al., 2011). La fuerza de trabajo estandarizó el Flujo Extensible de Eventos (*eXtensible Event Stream*, XES por sus siglas en inglés) como formato de intercambio de los registros de eventos en el año 2010.

Las técnicas de minería de procesos se emplean en múltiples sectores (van der Aalst et al., 2006; Pérez et al., 2012) como la industria, los servicios web, la inteligencia de negocios y la salud. En este último, son muy utilizadas estas técnicas por los beneficios que brindan a este sector tan crítico, entre los que se pueden encontrar determinar el tiempo total que demora la atención a un paciente y los servicios médicos que representan cuellos de botella que interrumpen el funcionamiento de otros procesos dentro de las instituciones sanitarias.

En el CESIM se desarrollan investigaciones sobre los beneficios que brinda la aplicación de la Minería de Procesos en el sector de la salud. Hasta ahora las investigaciones han estado dirigidas a la etapa de descubrimiento, teniendo como resultado el desarrollo de un procedimiento para detectar eventualidades mediante la aplicación de técnicas de descubrimiento. A nivel internacional entre los autores que han tratado la minería de procesos en el sector de la salud se encuentran (Yang & Hwang, 2006; Mans et al., 2008; Saravanan & Rama, 2010; van Doremalen, 2012).

1.2.1 Perspectivas para aplicar Minería de procesos

La minería de procesos puede abarcar diferentes perspectivas (van der Aalst, 2011; Yzquierdo, 2012; Leoni & van der Aalst, 2012; Saravanan & Rama, 2012) para realizar el análisis de los datos. Estas perspectivas describen el tipo de información que se puede extraer dependiendo de la información almacenada en el registro de eventos o de los objetivos que se persiguen con la aplicación de las técnicas de minería de procesos (Medeiros, 2006). A continuación se describen algunas de las perspectivas más utilizadas:

- **Perspectiva de control de flujo:** responde a la pregunta ¿Cómo son ejecutados en realidad los procesos? Se enfoca en determinar la estructuración del flujo de actividades del proceso de negocio (Yzquierdo, 2012). En esta dimensión se debe contar con información relativa a la identificación del caso o instancia ejecutada, la identificación de la actividad y el tiempo en el que se produjo la actividad.
- **Perspectiva de casos:** ¿Cuál es el rendimiento promedio para los casos de un determinado proceso?, puede ser una buena pregunta para guiar esta perspectiva. Además, se centra en las propiedades de los casos, teniendo en cuenta características como su ruta en el proceso, los actores que trabajan en él o los valores de los correspondientes elementos de datos (van Dongen, 2007). Se debe contar con información detallada de las actividades, como puede ser, la edad y el sexo del usuario involucrado.

- **Perspectiva organizacional:** responde a la pregunta ¿Quién realiza un trabajo determinado? trabajando sobre la información referente a los usuarios del sistema que ejecutaron cada actividad en específico. El objetivo es estructurar la organización al clasificar a las personas en términos de las funciones y roles, así como mostrar la relación entre los distintos usuarios (construir una red social) (Medeiros, 2006).
- **Perspectiva de tiempo:** se relaciona con la ocurrencia y frecuencia de los eventos, una pregunta frecuente que responder es ¿Cuánto demora en ejecutarse un proceso? Cuando los eventos tienen asociados marcas de tiempo, es posible descubrir cuellos de botella, medir niveles de servicio, monitorear la utilización de recursos y predecir el tiempo de procesamiento restante de casos en ejecución (van der Aalst et al., 2011).

En este trabajo se utiliza la perspectiva tiempo, ya que se pretende verificar la conformidad, en la duración y tiempo de espera de las actividades del procesos Solicitar producto en el módulo Almacén.

1.2.2 Registro de eventos

La mayoría de los sistemas de información modernos utilizan mecanismos para registrar la ejecución real de los procesos, poseen un registro de trazas, el cual es el punto de partida de la Minería de Procesos. Las técnicas de Minería de Procesos asumen que es posible registrar eventos secuencialmente (van der Aalst et al., 2011), donde cada proceso está compuesto por casos que no son más que instancias del mismo (Figura 1.2). Cada caso se compone de eventos que representan pasos bien definidos dentro del proceso; los casos contienen atributos o propiedades, donde los más usuales son la actividad que representan, la fecha y el usuario.

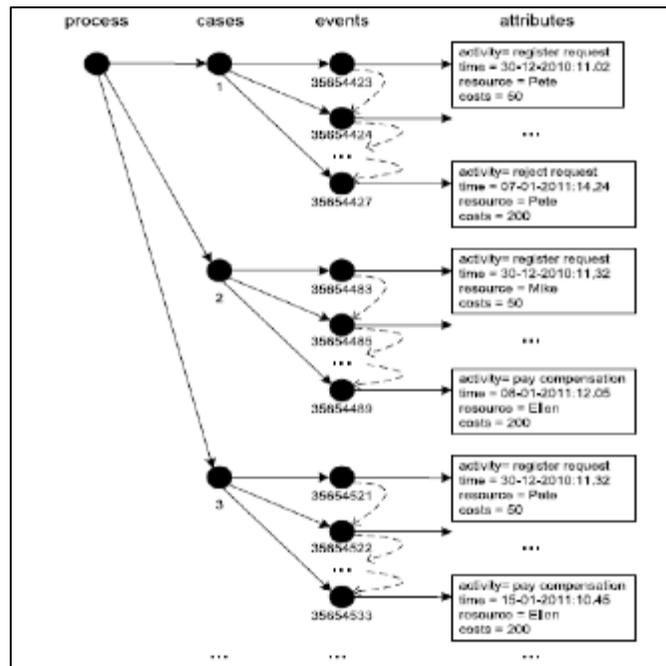


Figura 1.2. Resumen de un registro de eventos.

Fuente: van der Aalst, 2011

Los casos y cada evento en un registro utilizan un identificador mientras cada atributo de los eventos puede ayudar a extender el modelo con información extra. Dependiendo de la cantidad o tipo de información disponible será la perspectiva que se podrá asumir y la información que se podrá extraer. En la Figura 1.3 se expone un ejemplo de un registro de eventos donde se puede observar los identificadores y los atributos de los eventos.

| 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 | ... |
| 3 | 35654521 | 30-12-2010:14.32 | register request | Pete | 50 | ... |
| | 35654522 | 30-12-2010:15.06 | examine casually | Mike | 400 | ... |
| | 35654524 | 30-12-2010:16.34 | check ticket | Ellen | 100 | ... |
| | 35654525 | 06-01-2011:09.18 | decide | Sara | 200 | ... |
| | 35654526 | 06-01-2011:12.18 | reinitiate request | Sara | 200 | ... |
| | 35654527 | 06-01-2011:13.06 | examine thoroughly | Sean | 400 | ... |
| | 35654530 | 08-01-2011:11.43 | check ticket | Pete | 100 | ... |
| | 35654531 | 09-01-2011:09.55 | decide | Sara | 200 | ... |
| | 35654533 | 15-01-2011:10.45 | pay compensation | Ellen | 200 | ... |
| 4 | 35654611 | 06-01-2011:15.02 | register request | Pete | 50 | ... |
| | 35654643 | 07-01-2011:12.06 | check ticket | Mike | 100 | ... |
| | 35654614 | 08-01-2011:14.43 | examine thoroughly | Sean | 400 | ... |
| | 35654645 | 09-01-2011:12.02 | decide | Sara | 200 | ... |
| | 35654647 | 12-01-2011:15.44 | reject request | Ellen | 200 | ... |
| ... | ... | ... | ... | ... | ... | ... |

Figura 1.3. Ejemplo de un registro de eventos.

Fuente: van der Aalst, 2011

Según (van der Aalst et al., 2011) los Registros de eventos se pueden clasificar en cinco niveles en dependencia de su madurez (Tabla 1.1). Las técnicas de minería de procesos se le pueden aplicar a registros de niveles ***** (quinto nivel), **** (cuarto nivel) y *** (tercer nivel).

| Nivel | Caracterización |
|-------|--|
| ***** | Nivel más alto: el registro de eventos es de excelente calidad (confiable y completo) y los eventos están bien definidos. Los eventos se registran de manera automática, sistemática, confiable, y segura. Se toman en cuenta adecuadamente consideraciones acerca de la privacidad y la seguridad. Además los eventos registrados y todos sus atributos tienen una semántica clara. Esto implica la existencia de una o más ontologías. Los eventos y sus atributos se refieren a esta ontología. |

| | |
|------|---|
| **** | Los eventos se registran automáticamente y de manera sistemática y confiable. A diferencia de los sistemas operando a nivel ***, se da soporte de manera explícita a nociones tales como instancia de proceso y actividad |
| *** | Los eventos se registran automáticamente, pero no se sigue un enfoque sistemático para registrar los eventos. Sin embargo, a diferencia de los registros de eventos en el nivel **, hay algún nivel de garantía que los eventos registrados calzan con la realidad. Aunque se necesita extraer los eventos de una variedad de tablas, se puede asumir que la información es correcta. |
| ** | Los eventos se registran automáticamente como un subproducto de algún sistema de información. La cobertura varía y no se sigue un enfoque sistemático para decidir que eventos se registran. Además, es posible pasar por alto el sistema de información. Por lo tanto, pueden faltar eventos o estos pueden no registrarse correctamente. |
| * | Nivel más bajo: los registros de eventos son de mala calidad. Los eventos registrados pueden no corresponder a la realidad y pueden faltar eventos. Los registros de eventos en los cuales los eventos se registran manualmente suelen tener dichas características. |

Tabla 1.1 Niveles de madurez para los registros de eventos.

Fuente: van der Aalst et al., 2011

Los criterios para definir la calidad de estos registros abarcan características tan diversas como la confiabilidad, completitud, automatismo, sistematicidad, privacidad, seguridad y semántica clara en los datos almacenados. La calidad de las técnicas de Minería de procesos depende en gran medida de la calidad de estos registros, de ahí la importancia de trabajar con datos completos.

1.2.3 Clasificación de los procesos

Según (van der Aalst, 2011), los procesos se pueden clasificar en dos categorías: estructurados y desestructurados conocidos también como Proceso Lasaña y Espaguetis respectivamente. El mismo autor define que tomando una métrica de ajustes (*fitness*), esta tiene que tener al menos un valor de 0,8 en un Proceso Lasaña, que significa que más del 80% de los procesos se comportan tal como se había previsto. Estos procesos siguen patrones regulares, casi predictivos. En la Figura 1.4 se muestra un ejemplo de un Proceso Lasaña, se pone de evidencia lo estructurado del proceso. Los cuadros en blanco representan actividades que se relacionan siguiendo un patrón regular, mientras los cuadros en negro representan actividades invisibles que no afectan la predictibilidad del modelo.

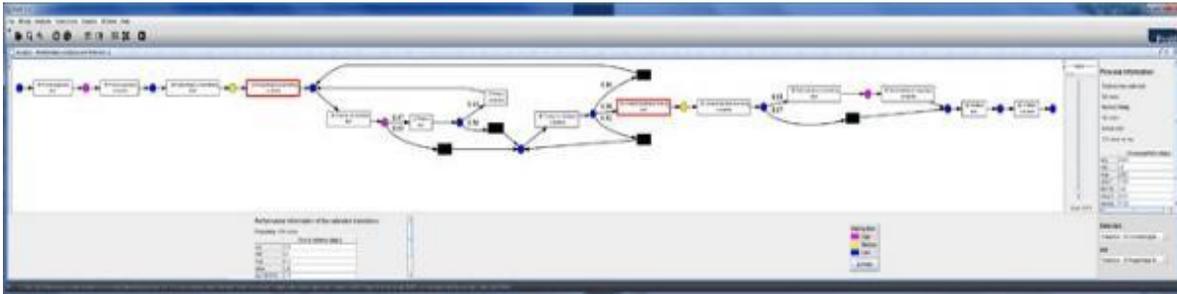


Figura 1.4. Ejemplo Proceso Lasaña.

Fuente: van der Aalst, 2011

Un Procesos Espaguetis o desestructurado posee muchas actividades relacionadas entre sí, lo cual dificulta hacer un análisis del modelo. En la Figura 1.5 se observa un ejemplo de lo complejo que puede ser realizar un análisis de un proceso con estas características.

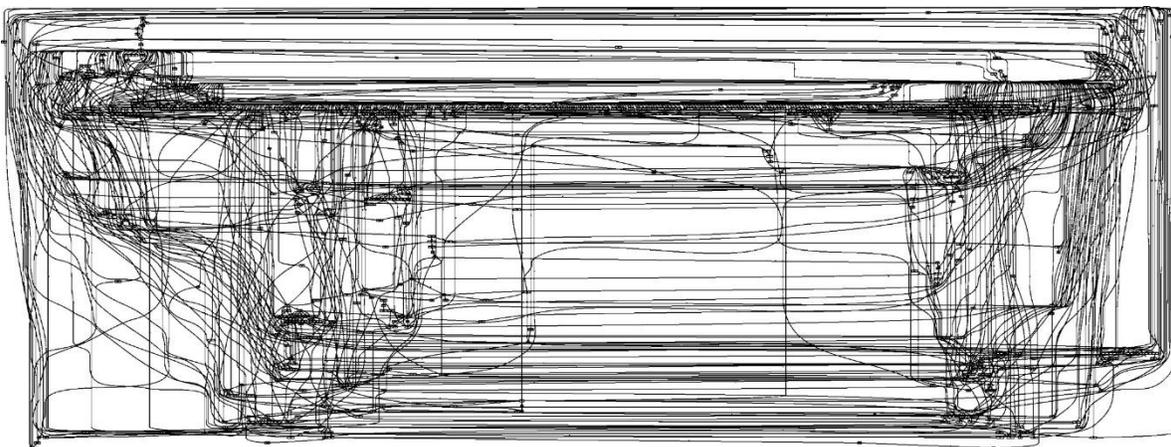


Figura 1.5. Ejemplo Proceso Espaguetis.

Fuente: van der Aalst, 2011

Además cuando el proceso es relativamente estructurado, el modelo de control de flujo puede ser extendido con otras perspectivas. Resumiendo, la clasificación de los procesos depende del orden en que se ejecuten las actividades y de las cantidades de estas relacionadas entre sí.

1.2.4 Técnicas de minería de procesos

Las técnicas son procedimientos o conjuntos de reglas que tienen como objetivo obtener un resultado determinado. Representan conocimiento transmitido de un individuo a otro, en forma de destrezas, como el uso de herramientas y saberes muy variados (Bozkaya, 2009).

Las técnicas en minería de procesos pueden variar en dependencia del tipo de investigación que se quiera realizar (van der Aalst et al., 2011). Específicamente las de descubrimiento son la base para realizar los demás tipos de minería, a continuación se describirán estas y seguidamente las técnicas utilizadas en chequeo de conformidad, ya que ambas serán usadas en la investigación.

Técnicas de descubrimiento

Se utilizan para identificar a partir del registro de evento un modelo representativo del proceso en el que queden reflejadas las dependencias entre las diferentes actividades que han sido registradas (Yzquierdo, 2012). Estas técnicas permiten identificar patrones que se manifiestan en los modelos. Algunos factores que afectan la calidad del modelo descubierto son el ruido y la desestructuración presente en el proceso analizado. A continuación se describen las técnicas de descubrimiento utilizadas en la investigación y las ventajas de su uso:

- Inspección de registro de eventos.

Para realizar la inspección es necesario visualizar el registro de eventos antes de aplicar alguna técnica de minería, con el fin de conocer la información almacenada en el registro. Solo se puede responder a determinadas preguntas si los datos están almacenados en dicho registro (Verbeek, 2010).

- Alineación de Trazas (*Trace Alignment*).

La alineación de trazas facilita el problema de determinar patrones interesantes en medianos y grandes registros, los que se determinan automáticamente y se muestran al usuario. Las trazas alineadas constituyen una representación de las tareas de acuerdo a un orden relativo y su estructuración en casos. El orden establecido entre las tareas permite identificar los patrones de flujo de control que se manifiestan en el proceso (Bose et al., 2012).

- Minería Difusa (*Fuzzy Miner*).

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. Según (Günther, 2009) está caracterizada por la presencia de dos tipos de nodos; nodos primitivos (se refieren a una tarea) y nodos que se refieren a un conjunto de tareas o clúster.

Además permite agrupar tareas, aunque considera 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, basados en la correlación entre las tareas y la importancia de una secuencia.

- Análisis de puntos de decisión (*Analyze using Dotted Chart*).

Esta técnica puede ser aplicada a la mayoría de los sistemas de información que generen registro de eventos, se utilizan principalmente para auditar y controlar procesos (Song & van der Aalst, 2007). Al analizarse adecuadamente un registro con esta técnica, se puede obtenerse un gran conocimiento lo que ayuda a las organizaciones a mejorar considerablemente la calidad de sus servicios.

Se enfoca en analizar cómo los atributos de datos influyen en las decisiones tomadas en el proceso sobre la base de las ejecuciones de procesos previos. Por tanto, se puede definir que su objetivo general es detectar las dependencias de datos que afectan la ruta de un caso. Está enfocado al aprendizaje automático y utiliza algoritmos implementados en Weka.

- Convertir árbol de procesos en una Red de Petri (*Convert Process Tree to Petri Net*)

Tras realizar el descubrimiento de un árbol de procesos con la técnica *Mine a Process Tree using Inductive Miner* a un registro de eventos seleccionado, se utiliza esta técnica que convierte dicho árbol en una Red de Petri. Esto permite posteriormente aplicar otras técnicas de minería de procesos al nuevo modelo generado por ejemplo, las técnicas que reproducen un registro utilizando una Red de Petri (replay) y que utilizan una Red de Petri como entrada.

- Adicionar eventos de inicio y fin artificiales (*Add artificial start and end events*)

Agrega a cada secuencia de actividades en el registro de trazas, eventos artificiales de inicio y de fin, con el objetivo de unir los diferentes flujos del proceso.

- Minar un árbol de procesos utilizando minería inductiva (*Mine a Process Tree using Inductive Miner*)

Esta técnica permite generar un árbol, donde los nodos representan actividades o transiciones. Las transiciones se representan con una flecha que apunta a la derecha, mientras los ciclos se representan con una x circundada por una flecha y las decisiones con una x sola. Este tipo de representación jerárquica es muy intuitiva y permite determinar con mayor facilidad la existencia de lazos y actividades finales. También permite seleccionar otra vista donde las actividades se representan con una notación

BPMN. Esta vista es especialmente útil para compárala con el diagrama de procesos de negocio original y detectar desviaciones en el comportamiento de los procesos.

Técnicas de chequeo de conformidad

Las técnicas de chequeo de conformidad pueden ser utilizadas para medir el rendimiento de los algoritmos de descubrimiento de procesos y reparar modelos que no se han ajustado a la realidad (van der Aalst, 2011). Comparan el comportamiento expresado en los modelos con el registrado en los registros de eventos y viceversa.

Estas técnicas se centran en dos aspectos: evaluar hasta qué punto un registro coincide con un modelo y destacar los puntos de discrepancia. Establecen un proceso de verificación de Inteligencia de Procesos de Negocio si ciertas propiedades se mantienen en un registro (Castellanos et al., 2009). A continuación se presentan las técnicas más importantes pertenecen a esta etapa de la minería de procesos:

- Verificación de conformidad utilizando *ETConformance*.

Al utilizar la técnica *ETConformance*, se puede medir la conformidad entre un registro y una Red de Petri (Muñoz-Gama, 2010), es decir, la calidad de la Red de Petri para representar el comportamiento descrito por el registro. En la versión actual, proporciona una métrica (PTET) para cuantificar la dimensión de precisión de la conformidad, es decir, la Red de Petri representa precisamente el registro, sin incluir el comportamiento adicional no reflejado en el registro.

Por otra parte, para ubicar el lugar donde se encuentran los problemas de precisión, automáticamente se envían las trazas de desconformidad mínimas (MDT), es decir, las trazas mínimas de tamaño aceptadas por la Red de Petri, pero no incluidas en el registro. En otras palabras, por cada traza $a_1 \dots a_n$ en MDT, $a_1 \dots a_{n-1}$ se encuentra en la Red de Petri y en el registro, pero $a_1 \dots a_n$ está solamente en la Red de Petri.

- Comprobador LTL (*LTL Checker*).

Esta técnica permite al usuario comprobar el registro con un modelo LTL determinado y especificando una serie de ajustes previamente, se utiliza mayormente para realizar auditorías. Existen dos versiones, la primera el comprobador LTL, la cual requiere un modelo LTL como entrada; la segunda

es el comprobador LTL predeterminado (*LTL Checker Default*), esta utiliza un modelo LTL predefinido y no requiere a ningún otro modelo LTL como entrada.

- Análisis de rendimiento con red de Petri (*Performance Analysis with Petri Net*).

Esta técnica proporciona un medio para evaluar el rendimiento de los procesos. Su principal objetivo es el de agregarle a los procesos, indicadores clave de rendimiento, para que puedan ser analizados de forma intuitiva. Para su uso debe recibir un registro de eventos y un modelo de procesos (en forma de una red de Petri) del proceso en que se está trabajando.

El modelo de proceso se utiliza como una interfaz, a través del cual la información de rendimiento se puede invocar. De esta información se puede obtener, por ejemplo las rutas XOR-splits y el tiempo por casos, representadas por fichas que ocurren en ciertos lugares dentro del proceso. Se distinguen cuatro grupos de información de rendimiento, tales como las métricas de proceso, las métricas de lugar, las métricas de dos transiciones y las métricas de actividad.

- Reproducir un registro utilizando una Red de Petri para el Análisis de Desempeño/Conformidad.

Recibe como entrada una Red de Petri y un registro de eventos que utiliza para generar alineaciones avanzadas entre cada una de las trazas existentes en el registro de eventos y la red descubierta. Para registrar un máximo beneficio al aplicar esta técnica deben satisfacerse los siguientes requisitos, primero la red debe tener asociado un marcador inicial, segundo debe poseer un marcador final, en la caso de no poseerlo la técnica asumirá que todos los estados en punto muerto/estados alcanzables en la red son marcas finales y por último el registro debe contener al menos una traza para analizar, no puede estar vacío.

1.2.5 Métricas de calidad

Los registros de trazas muchas veces presentan inconsciencias como ausencia de información y ruido, lo que a su vez provoca que los modelos descubiertos establezcan incorrectas relaciones en las actividades incorporadas a este (Yzquierdo, 2012.) Por lo cual es necesario medir el grado en el que el modelo descubierta representa el comportamiento observado en el registro de evento. Estas verificaciones se realizan al Chequeo de conformidad, para el cual han sido desarrolladas métricas para medir la calidad del mismo. A continuación se describen algunas de ellas (Chamorro, 2013):

- **Fitness:** Mide la capacidad de reproducir el registro de eventos (realizar replay). Un valor Fitness perfecto corresponde a 1, e implica que todas las trazas en el registro pueden ser reproducidas por el modelo de inicio a fin.
- **Simplicidad:** Se refiere a la Navaja de Occam (Chamorro, 2013), es decir, al modelo más simple que puede explicar el comportamiento que se ve en el registro. Por ejemplo, se puede medir por el número de nodos y arcos.
- **Precisión:** Un modelo es preciso si no permite muchos comportamientos diferentes, es decir, no es sub-ajustado al registro.
- **Generalización:** Un modelo es generalizado si no restringe el comportamiento que se aprecia en el registro, es decir, no es sobre-ajustado al registro.

La mayoría de las métricas mencionadas consideran que el modelo de proceso está representado por una red de Petri, aun cuando la métrica puede generalizarse para notaciones con semejante expresividad. Se sugiere considerar varias dimensiones de calidad, para obtener una apreciación más amplia del modelo.

1.3 Modelos para aplicar técnicas de minería de procesos

En el ámbito de Minería de procesos los modelos son equivalentes a las metodologías de desarrollo de software. Estos proponen una serie de pasos para preparar el registro de eventos y aplicarle las técnicas de minería de procesos.

A continuación serán descritos algunos de los principales modelos propuestos por figuras emblemáticas dentro del campo de la minería de procesos, con el objetivo de exponer el funcionamiento básico de los mismos, para realizar una selección de las etapas o flujos necesarios para dar respuesta a la problemática planteada en la presente investigación.

1.3.1 Modelo de Rozinat

Se desarrolla con el objetivo de extraer información clave (de los datos, el rendimiento y la organización), que puede ser usada para la creación de un modelo de simulación, que permite visualizar las dependencias entre las técnicas de minería de proceso a utilizar (Rozinat et al., 2009).

Paso 1: Creación del registro de eventos: La mayoría de los sistemas de información poseen algún tipo de registro de eventos, que contiene el flujo de actividades ejecutadas durante el proceso de negocio. Cada evento se refiere a una instancia de proceso y a una actividad, esta última se encuentra determinada por propiedades. El registro de eventos es transformado al formato MXML.

Paso 2: Descubrimiento del flujo de actividades: A partir de la información recopilada del registro de eventos, es empleado un algoritmo de descubrimiento que refleje las relaciones entre las actividades, por ejemplo el algoritmo *alpha* (α) que automáticamente construye el modelo en una red de Petri.

Paso 3: Análisis de puntos de decisión: Después del descubrimiento se pretende profundizar en la perspectiva de los datos del proceso para descubrir dependencias de datos que influyen en la ruta de un caso. Para analizar las opciones en un proceso, primero se debe identificar las partes del modelo en las que el proceso se divide en varias ramas alternativas, también llamadas puntos de decisión. Luego, se desea encontrar las reglas para el seguimiento de una ruta o la otra, sobre la base de los atributos asociados a los eventos en el registro.

Paso 4: Análisis de rendimiento: Se realiza con el objetivo de profundizar en la perspectiva del rendimiento del proceso, para mejorar el modelo del proceso con información sobre los tiempos de ejecución y tiempos de espera para las actividades. También se desea mejorar el modelo del proceso con las probabilidades para tomar caminos alternativos, y con información sobre el esquema de generación de caso.

Paso 5: Descubrimiento de roles: La minería organizacional pretende descubrir el modelo de la organización (las relaciones entre los recursos y sus roles o unidades funcionales) y las reglas de asignación (las relaciones entre los roles o unidades funcionales y actividades). Con sólo utilizar un registro de eventos, es difícil descubrir las diferencias entre estos conceptos. Por lo tanto, es necesario formar grupos de recursos en los que las personas realicen actividades similares, por ejemplo a partir del perfil.

Paso 6: Modelo integrado: En los pasos anteriores se evidencia cómo pueden ser extraídas de un registro de eventos diferentes características de un proceso. Estas pueden utilizarse para construir un modelo de simulación. Para obtener una mejor visión sobre el proceso en su conjunto, es útil integrar las perspectivas descubiertas en un modelo holístico. Esto es bastante fácil, siempre y cuando las

características del proceso descubiertas sean ortogonales entre sí (es decir, no hay información contradictoria). Si hay características conflictivas, entonces esto se convierte principalmente en un reto técnico.

1.3.2 Modelo de Bozkaya

Este modelo se desarrolla con el objetivo de ofrecer una visión general de los procesos dentro de los sistemas de información actuales en un corto período de tiempo (Bozkaya, 2009). Está compuesto por seis fases que se describen a continuación:

Fase 1: Preparación del registro de eventos: La mayoría de los sistemas de información poseen un formato personal de registro, el cual es necesario extraer y transformar. Primeramente se selecciona la mejor idea de un caso, luego se identifican las actividades y sus eventos, todo esto con el fin de obtener un registro de eventos adecuado para proceder a la siguiente fase.

Fase 2: Inspección del registro de eventos: El objetivo fundamental es entender con mayor claridad el registro de eventos y obtener una primera vista del proceso. Primeramente se hace un resumen estadístico del registro que da una idea de la cantidad de eventos almacenados y de los posibles algoritmos de minería a utilizar para obtener buenos resultados en las fases siguientes. Teniendo en cuenta las estadísticas, se filtra el registro de eventos con el objetivo de eliminar los casos incompletos; constituyendo este registro de eventos filtrado la entrada para las siguientes fases. Para obtener una primera vista, se usa la técnica *Fuzzy Miner*, con un evento de inicio y fin artificial, agregado a cada caso en el registro.

Fase 3: Análisis del flujo de actividades: Esta fase tiene como objetivo dar respuesta a la pregunta: "¿Cómo se ve el proceso actual?". Si se tiene una descripción del proceso, se ejecuta una verificación de conformidad para comprobar si el proceso se ajusta a la especificación, o sea, que cada caso en el registro de eventos puede ser reproducido en el proceso definido. Si no existe una descripción del proceso, el flujo de actividades necesita ser descubierto. Actualmente hay disponibles numerosos algoritmos para descubrir los procesos que derivan en modelos. Una buena comprobación para el modelo del proceso consiste en ejecutar una verificación de la conformidad sobre el mismo.

Fase 4: Análisis de rendimiento: A partir del descubrimiento del flujo de actividades del proceso, se pueden utilizar los modelos para analizar el rendimiento del proceso. Una de las preguntas que da respuesta a esta fase es: "¿Hay cuellos de botella en el proceso?".

Fase 5: Análisis de los roles: En caso de que el registro de eventos contenga información sobre quién ejecutó determinado evento, los roles en dicho proceso se pueden analizar. Durante el desarrollo de esta fase se responden preguntas como: "¿Quién ejecuta qué actividades?" y "¿Quiénes están trabajando juntos?".

Fase 6: Mostrar los resultados al cliente: Como resultado de las fases anteriores se obtiene el comportamiento real del sistema, que generalmente se desvía del proceso diseñado, puede ser tanto por conductas inadecuadas como por aportes al flujo de actividades para facilitar el trabajo. Este resultado debe ser tratado directamente con el jefe de la organización, quien determina cuáles fueron los comportamientos deseados y no deseados en el sistema. Luego utiliza todo este conocimiento para rediseñar su sistema de información para hacerlo más eficiente y competitivo.

1.3.3 Modelo de van der Aalst

Este modelo hace una descripción del ciclo de vida (L*) de un proyecto de minería de proceso (van der Aalst, 2011). Y está compuesto por cinco etapas que se describen a continuación:

Etapa 0 - Planificar y justificar. Los proyectos de minería de procesos comienzan por planificar y justificar las actividades a desarrollar. Para obtener adecuados resultados es necesario valorar los beneficios, antes de gastar esfuerzos en las actividades de minería. Existen tres tipos básicos de proyectos de minería de procesos:

- ✓ Basado en datos: impulsado por la disponibilidad de los datos de un evento; no posee preguntas concretas u objetivas, sino que espera surjan ideas valiosas al analizar los datos de eventos.
- ✓ Impulsado por preguntas: con el objetivo de responder a preguntas específicas, por ejemplo: ¿Por qué los casos atendidos por el equipo determinado tardan más que los casos atendidos por otro equipo?

- ✓ Impulsado por objetivos: con el objetivo de mejorar los procesos en relación a los indicadores claves de rendimiento (reducción de costos o tiempos de respuesta).

Etapa 1. Extracción. Luego de iniciar el proceso del proyecto se deben extraer del sistema los datos de eventos, modelos, objetivos y las preguntas. Para ello, es necesario entender los datos almacenados para analizar cuáles pueden ser utilizados en el análisis y qué preguntas son de mayor importancia.

Etapa 2. Crear el modelo de flujo de actividades y conectar con el registro de eventos. Tiene como objetivo fundamental obtener el modelo de flujo de actividades del proceso que se analiza. Dicho modelo puede ser descubierto usando las técnicas de descubrimiento de procesos; sin embargo, si es un buen modelo, se puede verificar a través del análisis de conformidad, comparar los modelos o combinar el modelo hecho a mano y el modelo descubierto.

Etapa 3. Crear un modelo de proceso integrado. La relación entre el registro de eventos y el modelo resultante de la Etapa 2 es utilizada para extender el modelo en esta etapa, el cual ha sufrido cambios factibles a través de la incorporación de nuevas perspectivas. A continuación se describen algunos pasos en el proceso de integración de las diferentes perspectivas:

Paso 3a: agregar la perspectiva organizacional. Es posible analizar la red social y posteriormente identificar las entidades organizativas que conectan las actividades de los grupos de recursos.

Paso 3b: agregar la perspectiva del tiempo. Las marcas de tiempo y las frecuencias se pueden utilizar para conocer las distribuciones de probabilidad que describen adecuadamente los tiempos de espera y servicio y las probabilidades de enrutamiento.

Paso 3c: agregar la perspectiva del caso. Los atributos incorporados en el registro son utilizados para la minería de decisión. Esto demuestra que los datos son relevantes y deben ser incluidos en el modelo.

Paso 3d: agregar otras perspectivas. Dependiendo de la información en el registro se pueden añadir otras perspectivas en el modelo como por ejemplo, la información sobre los riesgos y los costos.

Etapa 4. Apoyo a las operaciones. Esta etapa cumple con los objetivos de la minería de procesos, que son las actividades de apoyo operativo: detectar, predecir y recomendar. Se debe tener en cuenta que el apoyo operativo es la forma más ambiciosa de la minería de procesos. Esto sólo es posible para los procesos Lasaña. Por otra parte, es necesario que haya una infraestructura avanzada de las TIC que proporcione una alta calidad de los registros de eventos y permita la incorporación de un sistema de soporte operativo. Para poder alcanzar las Etapas 3 y 4 es necesario tener un proceso suficientemente estable y estructurado.

1.3.4 Valoración de los modelos anteriores

Luego de analizar cada uno de los modelos anteriormente descritos junto a los modelos de Jans (Jans, 2011) y Giessel (van Giessel, 2004) no detallados por su escaso aporte a esta investigación, se resumieron en la Tabla 1.2 las principales características que los definen. Como se puede observar los modelos más completos son el de van der Aalst y el de Bozkaya por lo que son los más aptos para utilizar en esta investigación.

Una vez consultadas las características de los modelos presentados, los autores de esta investigación consideran que por el aporte que brinda y su ajuste a las necesidades del trabajo el modelo a utilizar es el de van der Aalst. El cual da soporte a cualquiera de los tres tipos de minería de procesos, descubrimiento, conformidad y mejora.

Además no se enfoca en el uso de técnicas específicas, sino que permite aplicar una gran variedad de estas; las cuales a su vez posibilitan distintos puntos de vistas de los datos y realizar distintos análisis. Otra de las características que destacan a este modelo es que ofrece la oportunidad de realizar un análisis de los beneficios antes de iniciar el trabajo.

Funciona tanto en Procesos Lasaña (estructurado) como en Procesos espaguetis (desestructurado) y ofrece un marco bien definido para la aplicación de las técnicas de descubrimiento y conformidad.

| Características | Jans | Giessel | Rozinat | Bozkaya | van der Aalst |
|--|------|---------|---------|---------|---------------|
| Diagnóstico preliminar de los datos | | ✓ | | ✓ | ✓ |
| Extraer registro de eventos desde la fuente de datos | ✓ | ✓ | ✓ | ✓ | ✓ |

| | | | | | |
|--|---|---|---|---|---|
| Generar el flujo de actividades | ✓ | ✓ | ✓ | ✓ | ✓ |
| Caracterización del proceso de negocio | | | ✓ | ✓ | ✓ |
| Análisis a partir de tipo de proceso | | | | ✓ | ✓ |
| Obtener modelo integrado con la perspectiva tiempo | ✓ | | | | ✓ |
| Utilización de técnicas | ✓ | ✓ | ✓ | ✓ | ✓ |
| Análisis de rendimiento | | ✓ | ✓ | ✓ | ✓ |
| Análisis de roles | ✓ | ✓ | ✓ | ✓ | |
| Análisis con el cliente | | | | ✓ | ✓ |

Tabla 1.2 Comparación de modelos sobre minería de proceso

Fuente: Elaboración propia.

1.4 Herramientas

Se realizó un análisis de las principales tecnologías relacionadas con la Minería de Procesos y que se adaptaran a la presente investigación. Las herramientas a utilizar en la investigación son descritas a continuación.

1.4.1 Sistemas Gestores de Bases de Datos

Los Sistemas Gestores de Bases de Datos (SGBD) tienen la tarea de almacenar, modificar y extraer información de una base de datos. Además poseen mecanismos para insertar, editar y eliminar datos, así como proporciona métodos para mantener la integridad de los de los datos, para administrar el acceso de usuarios a los mismos y para recuperar la información si el sistema falla. La información contenida en los SGBD puede ser visualizada de varias formas, como informes o gráficos.

Entre los SGBD más utilizados en la actualidad se encuentran SQL Server (Yaofei, Huantong & Yinghua, 2011), propiedad de Microsoft, que destaca por su funcionalidad y fiabilidad e incorpora una herramienta para revelar información útil conocida como Big Data. Oracle Database, patentado por la empresa Oracle (Laszewski & Nauduri, 2012) es otro de los sistemas con más uso, destaca por las

facilidades que ofrece para desplegar y gestionar nubes de bases de datos, además de la compresión y organización por capas de datos, lo que permite densidades más altas para maximizar la eficiencia y flexibilidad de los recursos.

PostgreSQL es un potente SGBD, es de código abierto lo que posibilita utilizarlo, modificarlo y distribuirlo por todo el mundo de forma gratuita (Afyouni, Ray & Ilarri, 2013). También permite consultas complejas, utilizar disparadores y declara funciones. La versión 9.3.3 es la última estable del producto, que fue desarrollado por una comunidad conocida como *PostgreSQL Global Development Group* (PGDG). Por estas características, además de ser de amplia utilización en nuestra universidad, se decidió utilizar este sistema.

1.4.2 Clientes de administración de bases de datos

Los clientes de administración de bases de datos permiten gestionar las bases de datos de forma eficiente. Ofrecen una Interfaz Gráfica de Usuario (GUI del inglés *graphical user interface*) que agiliza y facilita el trabajo con los datos. Esta GUI facilita también la realización de consultas complejas y elaboración de grandes bases de datos.

EMS SQL Manager, es uno de los más utilizados debido a que permite añadir nuevas entradas e información, buscar datos dentro de los datos almacenados, agregar categorías y manipular objetos. Soporta algunas de las características avanzadas de PostgreSQL como espacios de tablas y argumentos nombrados.

Para la investigación se decidió utilizar la aplicación PgAdmin, ya que es el cliente más popular y rico en características de administración de código abierto y una plataforma de desarrollo para PostgreSQL, además que se ha convertido en los últimos tiempos en la herramienta más avanzada de código abierto en el mundo. Está diseñado para responder a las necesidades de todos los usuarios, como escribir consultas SQL sencillas para el desarrollo de bases de datos complejas. La última versión disponible en el mercado es PgAdmin III, este producto es desarrollado por una comunidad que está distribuida por todo el mundo. Esta característica que le permite ser actualizado constantemente dada su condición de software libre y está disponible en más de una docena de idiomas. Es distribuida bajo licencia PostgreSQL.

La interfaz gráfica es compatible con todas las características de PostgreSQL y facilita la administración. La aplicación también incluye un editor de resaltado de sintaxis SQL, un editor de código

del lado del servidor, un agente de planificación de tareas de SQL / batch / shell, el soporte para el motor de replicación Slonyl y mucho más. La conexión con el servidor se puede hacer a través de TCP / IP o Unix Domain (en plataformas Unix), y puede utilizar el encriptado SSL para preservar la seguridad. Además, tiene como ventaja que no requiere de controladores adicionales para comunicarse con el servidor de bases de datos.

1.4.3 Herramientas de Minería de Procesos

Las herramientas dedicadas a la Minería de Procesos permiten analizar Registros de eventos y extraer información útil y oculta de estos. Disco es una solución desarrollada y distribuida por Fluxicon (van der Aalst et al., 2009), que posee una interfaz muy intuitiva, que genera modelos animados. Las mediciones de rendimiento se muestran de una forma directa, además que puede lidiar con grandes registros de eventos y modelos complicados, sin dejar de hacer la conversión y el filtrado fáciles.

Para la investigación se seleccionó ProM, herramienta que implementa las técnicas de minería de procesos definidas previamente en esta investigación. Otra cualidad que define y justifica su selección es la rapidez y eficacia con que genera y modifica los modelos descubiertos por las técnicas de minería. Se compone de un conjunto de técnicas que permiten a la herramienta ajustarse a las necesidades de sus usuarios y de los tipos de minería de procesos que deseen utilizar. Permitiendo la integración y comprobación entre varios modelos o entre los registros de eventos y sus respectivos modelos.

Es una herramienta de código abierto especialmente diseñada para apoyar el desarrollo de la minería de procesos. Es desarrollada por Fluxicon y contiene una amplia variedad de técnicas en forma de plugins, la última versión del producto es la 6.3. Algunos de ellos van más allá de la minería de procesos (como hacer la verificación de procesos, la conversión entre diferentes notaciones de modelado, etc.). Para el trabajo con la herramienta se utilizan como entrada un registro de eventos solamente o un registro de eventos y un modelo de proceso (Verbeek, 2010).

Otra característica importante, es que es una plataforma independiente que se implementa en Java, y puede ser descargada sin tener que pagar por algún tipo de licencia para su uso. Además que ha sido utilizada y se han desarrollado plugin para ella en la universidad.

1.4.4 Herramientas para la extracción de registros de eventos

Las herramientas para la extracción de registros de eventos utilizan una fuente de datos para extraer estos registros que sirven como entrada a las herramientas de análisis de Minería de Procesos. Entre

las soluciones desarrolladas con este propósito se pueden citar Nitro (van der Aalst et al., 2011) y XESame (Verbeek, 2011), que son las más utilizadas en la actualidad (van der Aalst, 2011). Nitro destaca por su facilidad de uso y el rendimiento que ofrece ante los errores, inconsistencias y otros artefactos no deseados que aparecen en las fuentes de datos. En las últimas versiones se han optimizado los filtros de registros de forma tal que permiten obtener los registros de una forma rápida y productiva.

XESame, desarrollada por Fluxicon, proporciona una forma genérica para la extracción de un registro de eventos de alguna fuente de datos, y está diseñada para ser fácil de utilizar. Una fortaleza clave de XESame es que no se requieren conocimientos de programación. Toda la conversión de la fuente de datos de registro de eventos puede ser definida a través de la interfaz gráfica de usuario (Verbeek, 2011). Su última versión disponible es la v1.3 incorporada a la herramienta Prom v6.3.

Esta herramienta fue diseñada con el objetivo de poder ser utilizada por cualquier persona aún sin ser un experto en el dominio de la programación en java. Todas estas características influyeron en que se seleccionara para ser utilizada en la investigación, además de que está incluido en el tutorial del ProM v6 y su última versión siempre está disponible en el sitio oficial del ProM.

En nuestro centro el pasado año se desarrolló la herramienta JyX v1.0, que permite facilitar la conversión de registro de eventos a archivo de extensión .XES que puede ser utilizado por la herramienta ProM. Su funcionamiento se basa en estructurar jerárquicamente los datos contenidos en el registro de eventos para simplificar su posterior escritura en un archivo. Existen dos versiones de esta herramienta que permiten utilizarla tanto en la plataforma de Linux como Windows.

Conclusiones del capítulo

1. Se explicaron los conceptos necesarios, para comprender el campo de acción de la investigación, y los elementos relacionados a esta.
2. Se describieron y analizaron los modelos existentes para aplicar técnicas de minería de procesos y se concluyó utilizar el modelo de van der Aalst por las características que posee.
3. Con el análisis de la bibliografía, se puso de manifiesto la carencia de investigaciones sobre chequeo de conformidad en sistemas hospitalarios nacionales y se referenciaron aquellas publicaciones de mayor aporte científico a la investigación en el marco internacional.
4. Se analizaron las técnicas de descubrimiento y conformidad más idóneas para su aplicación en la investigación.
5. El estudio de la disciplina puso de manifiesto las potencialidades de su empleo en diferentes sectores de la sociedad.

CAPÍTULO 2. PROCEDIMIENTO PARA EL CHEQUEO DE LA CONFORMIDAD APLICANDO TÉCNICAS DE MINERÍA DE PROCESOS.

En este capítulo se describe el modelo de van der Aalst y sus etapas, para realizar primeramente el descubrimiento y luego el chequeo de conformidad. En cada etapa se explican las fases que la componen y las tareas que se ejecutan.

2.1 Características del modelo de van der Aalst

El modelo de van der Aalst tiene como objetivo ofrecer un marco para poder aplicar las técnicas y herramientas utilizadas en la minería de procesos (van der Aalst et al., 2011). Una de las ventajas de su uso y de las características que lo han hecho tan popular es la claridad y cantidad de información disponible, además de casos de estudio y proyectos exitosos donde ha sido utilizado. Es aplicable a proyectos que abarquen a cualquiera de los tres tipos de minería de procesos, descubrimiento, conformidad y mejora. En la Figura 2.1 se describe el ciclo de vida de este modelo.

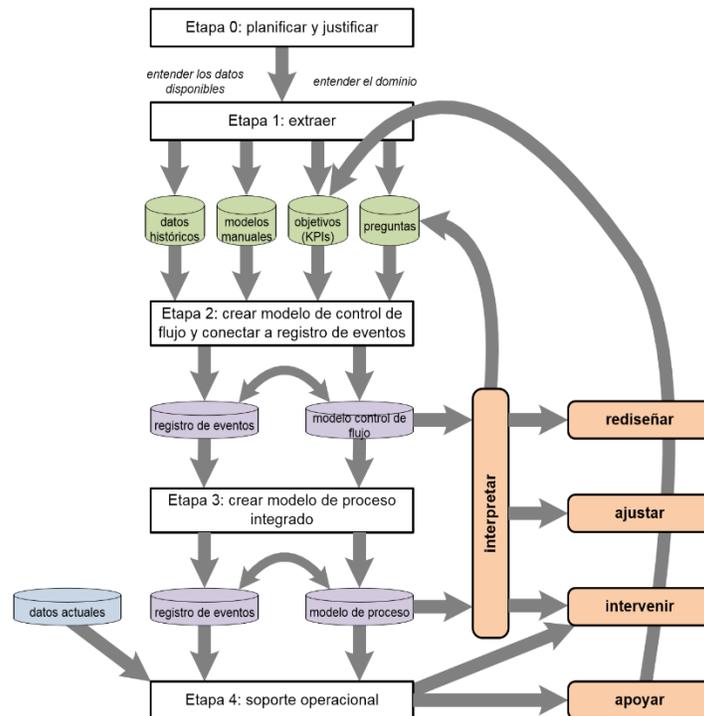


Figura 2.1. Ciclo de vida del modelo de van der Aalst.

Fuente: van der Aalst, 2011.

2.2 Descripción de las etapas y fases del modelo de van der Aalst

A continuación se describe de forma teórica cómo se estructura el modelo de van der Aalst, para realizar el chequeo de conformidad al proceso Solicitar producto del módulo Almacén.

2.2.1 Etapa 0 - Planificar y justificar.

Objetivo: Identificar y analizar el tipo de información a utilizar para generar el registro de eventos que represente el funcionamiento real de la aplicación.

En esta etapa se realizó un análisis de los módulos del HIS que cumplieran los criterios necesarios para aplicarle las técnicas de conformidad (van der Aalst, 2011), estos son: contar con suficiente información estructurada, de forma tal que permita modelar el comportamiento del sistema y utilizar el motor de flujo de trabajo JBPM. Esta última condición fue decisiva para la selección del módulo Almacén, ya que era el único módulo que contenía descripciones BPMN.

Para la correcta ejecución de un proyecto de minería de procesos, es necesario seleccionar acertadamente los objetivos que se persigue con el mismo (van der Aalst, 2011). En la presente investigación se siguió un enfoque basado en datos, por lo que se le aplicaron al registro de eventos múltiples técnicas de descubrimiento y conformidad, sin tener en cuenta objetivos o preguntas.

Las ventajas de este enfoque están en la profundidad del análisis que propicia ya que no está influenciado por información a priori. Además que aprovecha todas las oportunidades que ofrece la disponibilidad de datos y no es parcial con un punto de vista en específico.

2.2.2 Etapa 1 - Extracción.

Objetivo: Extraer del registro de trazas del sistema los datos necesarios para generar el registro de eventos.

A esta etapa se le adaptaron dos fases que permitirán obtener con un grado mayor de eficiencia los resultados de la investigación. A continuación se describen las dos fases a ejecutar:

Fase 1. Localizar y describir la fuente de datos.

Objetivo: Seleccionar el proceso a analizar y realizar un estudio de sus datos con el fin de obtener información básica del sistema.

En la Figura 2.2 se muestra en que tablas se localizan los datos necesarios para poder generar el registro de eventos. De la tabla `jbpm_processdefinition`, se extrae el identificador del proceso seleccionado y el atributo `name`. En la segunda tabla `jbpm_processinstance` se selecciona el identificador de las instancias de proceso perteneciente al proceso analizado. En la tabla `jbpm_taskinstance` se selecciona el identificador de las instancias de las tareas realizadas por los usuarios y los atributos: `name`, `actorid` y `create`. Para finalizar en la tabla `jbpm_pooledactor` solamente se utilizará el atributo `actorid` que representa a los departamentos a los que pertenecen los usuarios que interactúan con la aplicación.

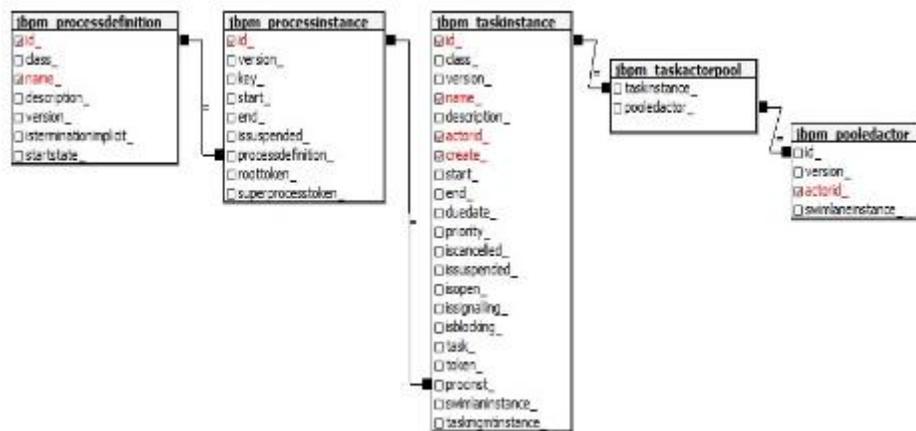


Figura 2.2. Localización del registro de trazas.

Fuente: Elaboración propia.

Fase 2. Generar el registro de eventos.

Objetivo: Utilizar la información obtenida en la fase anterior para crear el registro de eventos.

En esta fase se realiza la tarea de generar el registro de eventos y a continuación se mostrará según (Harmen Roest, 2012, van der Aalst & Weijters, 2004, van Dongen & van der Aalst, 2005) cuáles son los requisitos mínimos que debe tener un registro de eventos:

- ❖ Los eventos se corresponden con las actividades en el proceso de negocio.
- ❖ Todos los eventos están asociados con un caso particular.
- ❖ El orden temporal de las actividades debe ser representado (por ejemplo, mediante las marcas de tiempo).

En esta fase se propone seguir los siguientes pasos para que el registro de eventos resultante sea lo más ajustado posible al proceso que se está analizando.

Objetivo: Extraer desde la fuente de datos las trazas que se utilizarán para generar el registro de eventos.

A continuación se describen los pasos para crear el registro de eventos utilizando la herramienta XESame 1.3.

Herramienta XESame 1.3

Permite extraer directamente desde las bases de datos en que se encuentran registradas, las trazas del proceso y analizar los datos necesarios para generar el registro de eventos. Su objetivo principal es recolectar estas trazas y estructurarlas jerárquicamente para luego proceder a organizarlas en un fichero de extensión *.xes que luego será importado por la herramienta ProM.

Funcionamiento de XESame

A continuación se presentará la secuencia de pasos que se debe seguir para generar el registro de eventos:

Primeramente se procede a configurar la conexión de la herramienta con la base de datos, lo que permite extraer directamente desde cada tabla los atributos que conformaran el registro de eventos. En la Figura 2.3 se muestra la interfaz que visualizará el usuario.

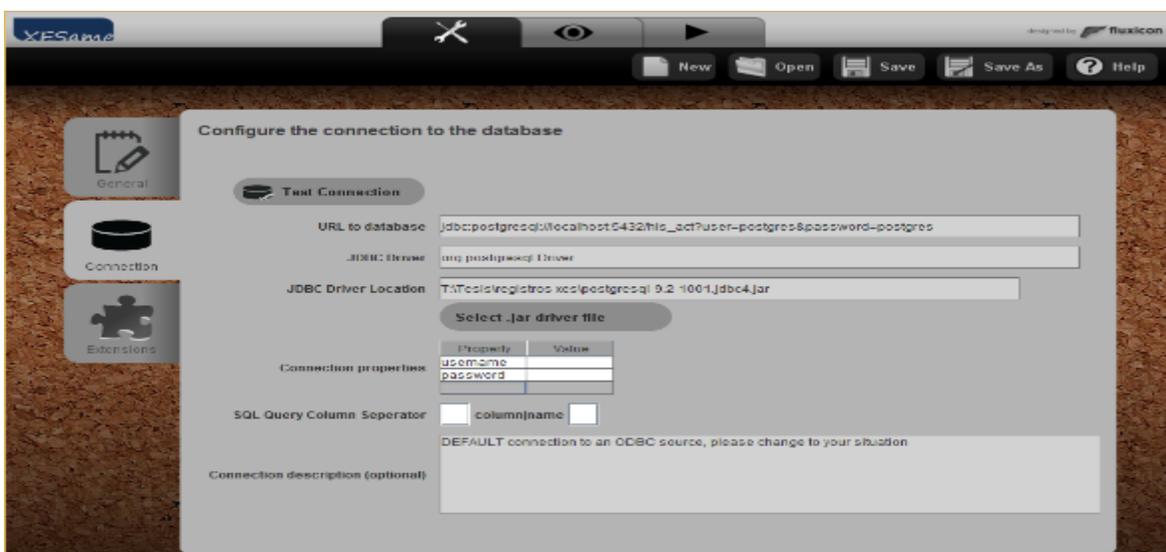


Figura 2.3. Interfaz de configuración para la conexión.

Fuente: Elaboración propia.

En la Figura 2.4, se observa cómo debe ser configurada la etiqueta "Registro" que es la base del archivo de extensión .xes en que estará contenido el registro de eventos. Una vez seleccionada esta etiqueta se debe acceder a la pestaña "Properties", aquí solamente se especificará de qué tabla de la base de datos se extraerán los datos relevantes para conformar el registro de eventos mediante la fila "From".

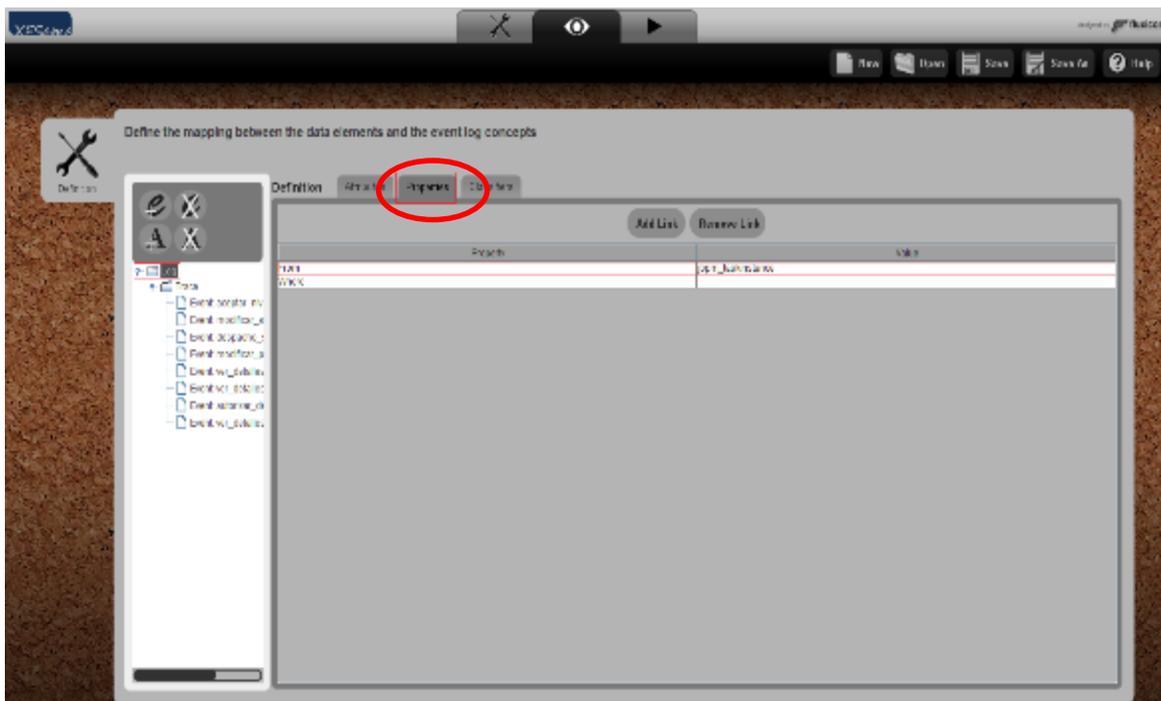


Figura 2.4. Interfaz de configuración para la etiqueta Registro.

Fuente: Elaboración propia.

En la Figura 2.5, se observa cómo queda configurada la etiqueta "Trace" y al igual que en la etiqueta anterior aquí también se debe trabajar sobre la pestaña "Properties". En la fila "From" se especifica la tabla de la base de datos en que se localiza la información a extraer. En fila "Where" se define la condición para la selección, es decir, si se filtra por nombre, id o algún otro atributo específico. En el caso de esta investigación se utilizó el filtrar por nombre cada una de las actividades que pertenecieran a las instancias del proceso Solicitar producto. En la etiqueta "TraceID" se define en qué columna de la tabla seleccionada en la base de datos se localiza el identificador de las instancias del proceso.

Luego las etiquetas "Link" se utilizan para conectar las tablas en la base de datos en caso de que se necesite algún atributo que no se encuentre en la tabla seleccionada inicialmente en la fila "From". Un ejemplo de este caso es si se observa la fila "Where", en esta fila el valor que toma pertenece a la tabla *jbpm_processdefinition*, debido a esto se generan dos filas "Link" que permiten realizar la unión de las tablas *jbpm_taskinstance* con *jbpm_processinstance* y esta última con *jbpm_processdefinition*, para que pueda ser seleccionado el nombre del proceso que se va a extraer para generar el registro de eventos.

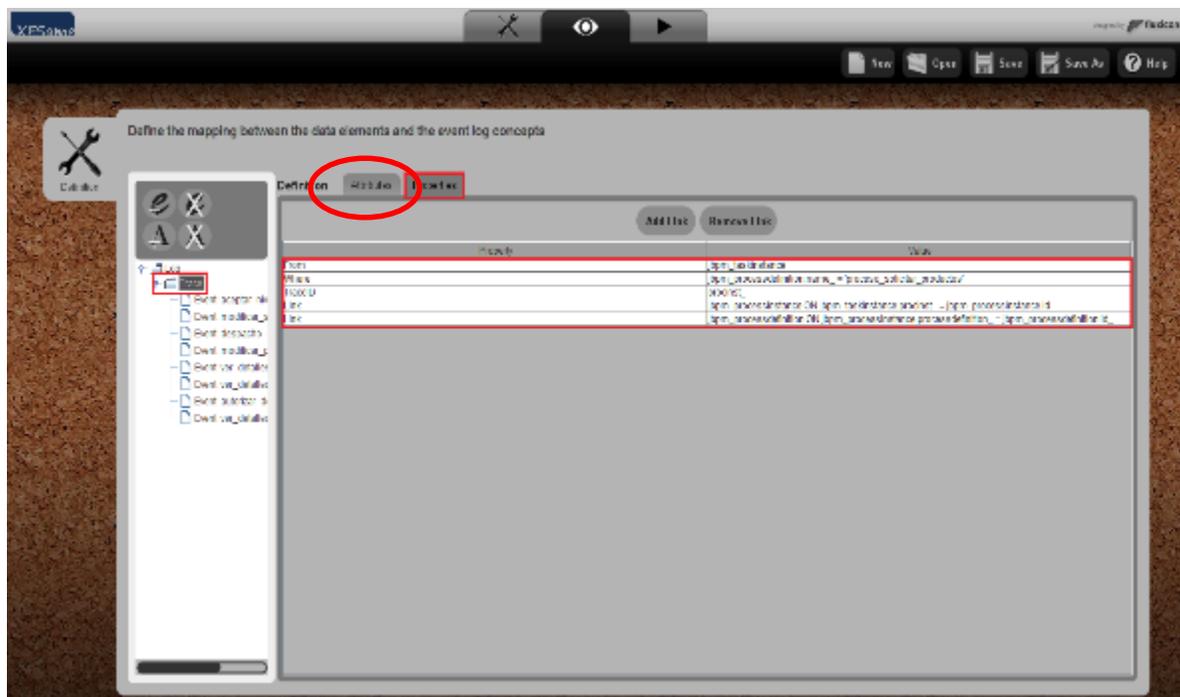


Figura 2.5. Interfaz de configuración para la etiqueta *Trace*.

Fuente: Elaboración propia.

A continuación se muestra cómo configurar en la etiqueta "Event" para que quede conformado correctamente el registro de eventos. En la pestaña "Attribute" se definen las etiquetas que compondrán a cada evento, por ejemplo:

- ✓ concept:instance
- ✓ concept:name

- ✓ lifecycle:transition
- ✓ org:resource
- ✓ org:group
- ✓ org:role
- ✓ time:timestamp
- ✓ semantic:modelReference

El valor que toman estos atributos se especifica en la columna "Value" y deben especificar la tabla de la que son extraídos y luego el nombre de la columna en la que se encuentran en esa tabla. Lo anteriormente descrito se muestra en la Figura 2.6.



Figura 2.6. Interfaz de configuración para la etiqueta *Event*.

Fuente: Elaboración propia.

En la pestaña "Properties" de la etiqueta "Event", se configura la consulta a la base de datos de la misma forma que en las "Properties" de la etiqueta "Trace" (Figura 2.7).

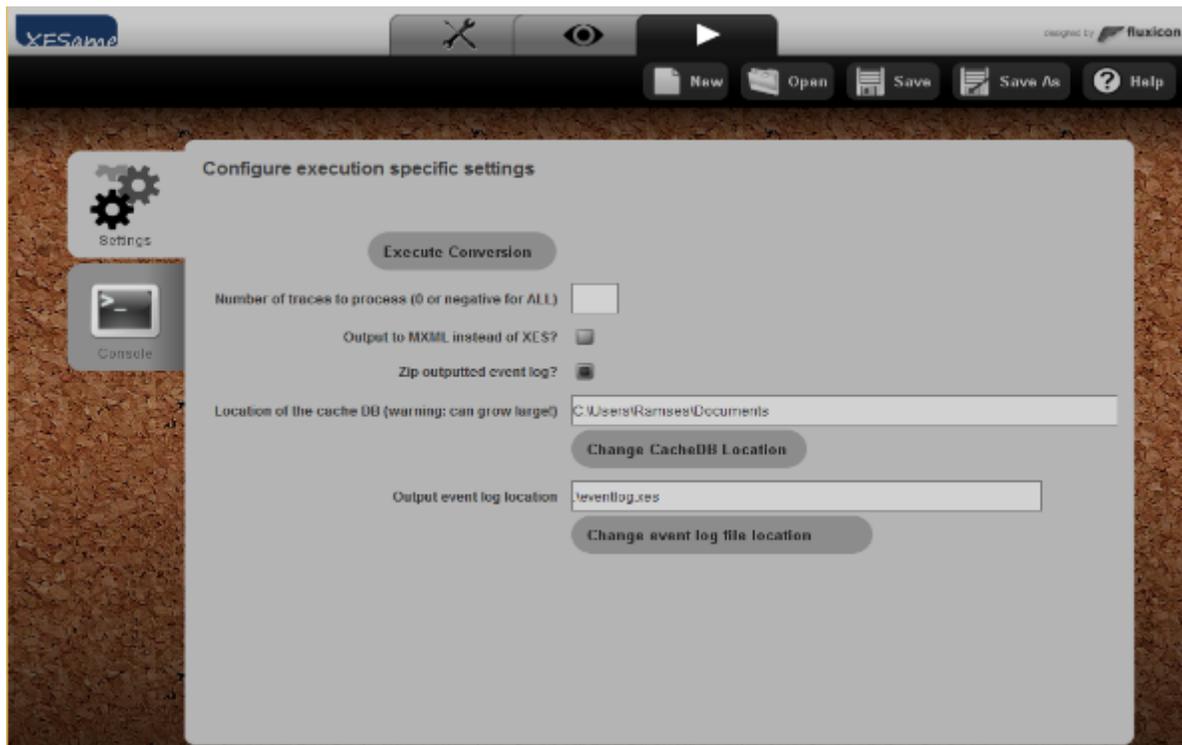


Figura 2.8. Interfaz que convierte el registro de eventos a fichero *.xes.

Fuente: Elaboración propia.

2.2.3 Etapa 2 - Crear el modelo de flujo de actividades y conectar con el registro de eventos.

Objetivo: Generar un modelo de procesos utilizando el registro de eventos obtenido en la etapa 1 y mostrar en dicho modelo el flujo de las actividades del proceso en análisis.

A continuación se muestran las fases a seguir para obtener el modelo preliminar del sistema:

Fase 1. Inspección del registro de eventos.

Objetivo: Identificar la cantidad de procesos, casos, eventos, clases de eventos y los originadores, así como la descripción de cada una de las instancias de proceso que se encuentra en el registro.

En esta fase comienzan a aplicarse las técnicas de minería de procesos que contiene la herramienta ProM 6.3. Para visualizar las características del registro de eventos se utiliza la opción "View" y dentro de esta la pestaña "Dashboard" (Figura 2.9). Finalmente el registro de eventos queda conformado por

Además posibilita explorar visualmente el flujo de actividades con sus respectivos atributos y por último permite observar los atributos del registro de eventos, lo anteriormente descrito se muestra en la Figura 2.11.



Figura 2.11. Flujo de actividades de cada instancia de proceso.

Fuente: Elaboración propia.

La opción "View" también posee la pestaña Summary (Figura 2.12), que permite visualizar un resumen detallado referente al registro de eventos, como ocurrencia absoluta y relativa de las actividades.

| Total number of classes: 8 | | |
|---|------------------------|------------------------|
| Class | Occurrences (absolute) | Occurrences (relative) |
| ver_detalle_sol_almacen+standard | 1172 | 48,833% |
| autorizar_despacho+standard | 1128 | 47,0% |
| ver_detalle_sol_bq+standard | 36 | 1,5% |
| despacho_sol_bq+standard | 36 | 1,5% |
| modificar_pedido_bq+standard | 15 | 0,625% |
| aceptar_niveles+standard | 7 | 0,292% |
| ver_detalle_sol_licitacion+standard | 4 | 0,167% |
| modificar_solicitud_licitacion+standard | 2 | 0,083% |
| Start events | | |
| Total number of classes: 3 | | |
| Class | Occurrences (absolute) | Occurrences (relative) |
| ver_detalle_sol_almacen+standard | 1172 | 96,7% |
| ver_detalle_sol_bq+standard | 36 | 2,97% |
| ver_detalle_sol_licitacion+standard | 4 | 0,33% |
| End events | | |
| Total number of classes: 6 | | |
| Class | Occurrences (absolute) | Occurrences (relative) |
| autorizar_despacho+standard | 1128 | 93,069% |
| ver_detalle_sol_almacen+standard | 44 | 3,63% |
| despacho_sol_bq+standard | 21 | 1,733% |
| modificar_pedido_bq+standard | 14 | 1,155% |
| aceptar_niveles+standard | 4 | 0,33% |
| ver_detalle_sol_bq+standard | 1 | 0,083% |

Figura 2.12. Probabilidades absolutas y relativas de ocurrencia de las actividades.

Fuente: Elaboración propia.

El análisis de las técnicas anteriores permitió identificar tres flujos de actividades bien definidos, sin embargo algunas instancias finalizan en actividades que no completan el proceso Solicitar producto. Esta inconsistencia se eliminaría si los distintos caminos que sigue el proceso tuvieran una actividad de inicio y de fin común. También permitió identificar que el flujo principal del proceso compuesto por las actividades `ver_detalle_sol_almacen` y `autorizar_despacho` abarca más del 95% de las ejecuciones de actividades.

Fase 2. Descubrir y obtener el modelo de procesos.

Objetivo: Aplicar la etapa descubrimiento utilizando varias técnicas de análisis de eventos de la minería de procesos.

En esta fase se representa el flujo de actividades del proceso analizado, para ello se utilizaron varias técnicas, a continuación se muestra el orden en que se deben utilizar estas para obtener un resultado satisfactorio. Un primer grupo de técnicas necesitan que se le agreguen actividades artificiales de inicio y fin para unir los distintos flujos que componen al proceso Solicitar producto y no se vean cada uno de estos flujos por separado:

1. “Add artificial start and end events” recibe como entrada el registro de eventos generado con la herramienta XESame y le agrega a este actividades artificiales de inicio y de fin.
2. “Minería Difusa (Fuzzy Miner)” recibe como entrada un registro de eventos y genera un modelo de alto nivel donde se representan los lazos y flujos de actividades.
3. “Alineación de Trazas (Trace Alignment)” recibe un registro de eventos y genera un modelo donde se pueden determinar los distintos flujos que sigue el proceso.

Al utilizar las técnicas que se describen a continuación no es necesario agregarle las actividades de inicio y fin artificiales, ya que estas le añaden un marcador al inicio y otro de fin, para unir los flujos y de esta forma evitar añadir actividades ficticias:

1. “Mine a Process Tree using Inductive Miner” recibe como entrada un registro de eventos y genera un árbol de procesos que representa mediante un modelo BPMN al flujo de las actividades del procesos bajo análisis.
2. “Convert Process Tree to Petri Net” recibe como entrada un árbol de procesos y lo transforma en una red de Petri equivalente.

2.2.4 Etapa 3 - Crear un modelo de proceso integrado.

Objetivo: Obtener un modelo de procesos al que se le haya incorporado un nuevo enfoque a partir de una perspectiva de la minería de procesos.

En esta fase se aplica la técnica *"Replay a Registro on Petri Net for Performance/Conformance"*, la cual posibilita detectar las desviaciones en el tiempo de cumplimiento de las diferentes solicitudes de productos. Está compuesta por tres fases que se describirán a continuación.

Fase 1. Selección de la técnica.

Objetivo: Seleccionar una técnica de chequeo de conformidad que se ajuste eficientemente a los atributos y características del proceso.

Se seleccionó la técnica *"Replay a Registro on Petri Net for Performance/Conformance"* por estar directamente relacionada con el análisis temporal. Además mediante una escala de colores muy intuitiva, clasifica las actividades por colores según su desviación con respecto al tiempo medio de estancia de las transiciones.

Fase 2. Obtener el modelo integrado enfocado a la perspectiva tiempo.

Objetivo: Obtener un modelo orientado a la perspectiva tiempo que permita realizar el chequeo de conformidad entre el modelo del sistema y el modelo integrado descubierto.

Fase 3. Realizar el chequeo de conformidad al modelo integrado.

Objetivo: Comprobar la aproximación existente entre el modelo del sistema y el modelo integrado descubierto.

2.2.5 Etapa 4 - Apoyo a las operaciones.

Objetivos: Analizar el modelo integrado en busca de funciones anómalas que presente el sistema con el fin de detectar, predecir y monitorear los errores que estén afectando o puedan afectar el buen funcionamiento de la aplicación.

En esta etapa se utilizará la técnica *"Analyze using Dotted Chart"*, esta permitirá seguir una línea temporal y realizar un análisis profundo del comportamiento de cada una de las instancias del proceso Solicitar producto almacenadas en el registro de eventos. Además de identificar los patrones de

comportamiento que presentan cada una de las instancias y en caso de existir desviaciones pueden ser vistas y detectadas por el usuario que esté aplicando este análisis.

Conclusiones del capítulo

1. Se describieron y analizaron las etapas que componen el modelo de van der Aalst.
2. Se mencionaron los objetivos que se persiguen en cada una de las fases que componen las etapas.

CAPÍTULO 3: APLICACIÓN DEL PROCEDIMIENTO DE CONFORMIDAD AL MÓDULO ALMACÉN

En este capítulo se realiza la validación de la aplicación del procedimiento de conformidad al módulo Almacén, a través de la puesta en práctica de las técnicas de descubrimiento y conformidad.

3.1 Módulo Almacén del Sistema de Información Hospitalaria

El Sistema de Información Hospitalaria es un sistema de gestión que permite a los hospitales la recolección, almacenamiento, procesamiento, recuperación y comunicación de información de atención al paciente y administrativa para todas las actividades relacionadas con el hospital. El mismo está concebido para llevar el control de las actividades de salud orientadas a los pacientes, permitiendo además gestionar y controlar los recursos de cada una de las áreas de las instituciones hospitalarias.

El módulo Almacén se encarga de gestionar el flujo de información sobre los distintos movimientos que puede tener un producto en un almacén. Existen tres tipos de solicitudes de producto, solicitud de licitación, solicitud de almacén y solicitud de bloque quirúrgico. A partir de estas actividades se desprenden el resto como se observa en la Figura 3.1.

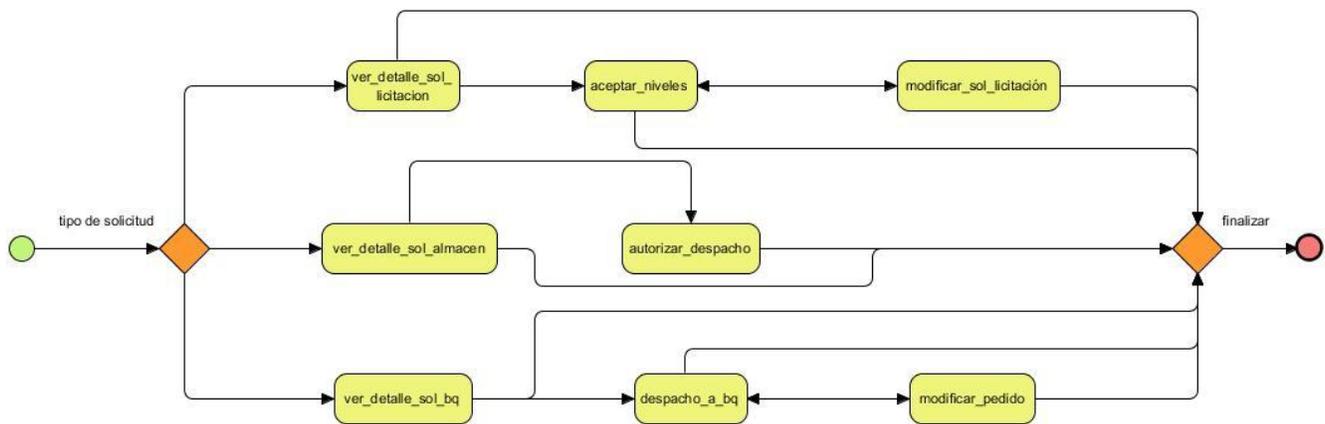


Figura 3.1. Diagrama de Procesos del Negocio proceso “Solicitar producto”.

Fuente: Elaboración propia.

3.2 Aplicación del procedimiento para el chequeo de la conformidad

3.2.1 Etapa 0 - Planificar y justificar

Para desarrollar esta investigación se desea detectar las irregularidades existentes en el módulo Almacén del Sistema de Información Hospitalaria. En este proceso se tuvo en cuenta la realización de un enfoque basado en datos, para realizar un análisis de conformidad al proceso Solicitar producto y la obtención de información proveniente de distintos modelos, enfocados en mostrar los datos contenidos en el registro de eventos desde diferentes puntos de vistas, para poder aplicar las técnicas de conformidad.

3.2.2 Etapa 1 - Extracción

Fase 1. Localizar y describir la fuente de datos

La base de datos se compone por tablas (Figura 3.2) que permiten apreciar las relaciones existentes en el sistema y tiene como objetivo registrar las acciones del usuario cada vez que acceda a su sección. Los datos que conforman el registro de trazas se contienen en cuatro tablas que registran cada una de las actividades realizadas por el usuario una vez haya iniciado su sección.



Figura 3.2. Schema public, donde se encuentra almacenada la información del módulo Almacén.

Fuente: Elaboración propia.

Fase 2. Generar el registro de eventos

Una vez realizado el estudio de la fase anterior se procede a acceder mediante la herramienta XESame a la base de datos para extraer los datos necesarios para generar el registro de eventos. Esta herramienta permite realizar consultas a la base de datos, permitiendo extraer directamente desde esta cada una de las propiedades necesarias para poder generar el registro de eventos. En la Figura 3.3 se muestra como queda la configuración de las propiedades de las trazas que conforman el registro de eventos.

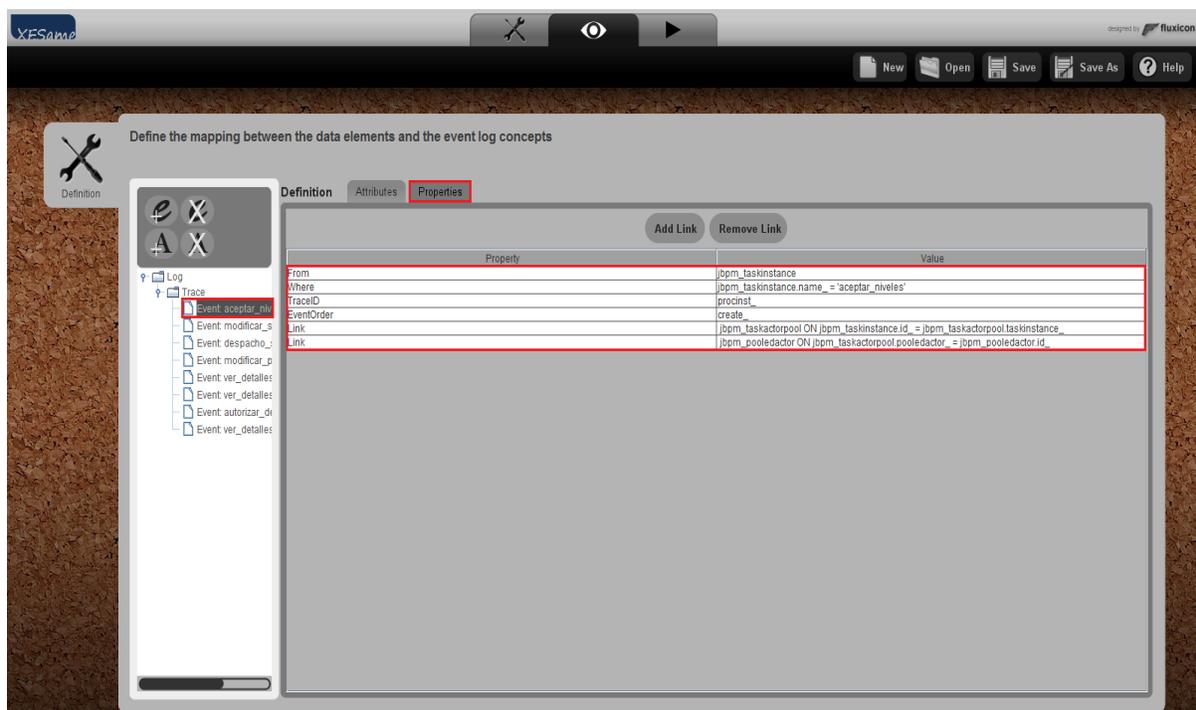


Figura 3.3. Consulta realizada por la herramienta XESame a la base de datos.

Fuente: Elaboración propia.

Tras realizar este procedimiento de extracción de datos la herramienta XESame exporta un fichero de extensión .xes que es el que contiene el registro de evento que permitirá aplicar las técnicas y

algoritmos de minería de procesos seleccionados para dar solución a esta investigación. En la Figura 3.4 se muestra un fragmento del archivo resultante de esta transformación.

```

<trace>
  <string key="concept:name" value="100564"/>
  <event>
    <string key="concept:instance" value="100566"/>
    <string key="concept:name" value="ver_detalle_sol_almacen"/>
    <string key="org:group" value="almacen_c_almacen_1"/>
    <string key="lifecycle:transition" value="standard"/>
    <date key="time:timestamp" value="2013-06-03T07:52:50.382+02:00"/>
  </event>
  <event>
    <string key="concept:instance" value="100597"/>
    <string key="org:resource" value="pepe@uci.cu"/>
    <string key="concept:name" value="autorizar_despacho"/>
    <string key="org:group" value="almacen_cc_almacen_14"/>
    <string key="lifecycle:transition" value="standard"/>
    <date key="time:timestamp" value="2013-06-03T07:52:57.677+02:00"/>
  </event>
</trace>

```

Figura 3.4. Muestra del registro de eventos obtenido con la herramienta XESame

Fuente: Elaboración propia.

La Figura 3.4 también representa una instancia perteneciente al proceso del sistema Solicitar producto, en ella se detallan los atributos que conforman a cada traza y las tareas porque está conformada.

3.2.3 Etapa 2 - Crear el modelo de flujo de actividades y conectar con el registro de eventos.

Una vez obtenido el registro de eventos se procede a utilizar la herramienta ProM para generar el modelo de flujo de actividades del Proceso Solicitar producto. Varios son los modelos que se pueden generar en esta etapa, cada uno de ellos permite una vista diferente del flujo de actividades.

Primeramente se le agregaron las actividades de inicio y fin al registro de eventos, utilizando la técnica *"Add artificial start and end events"*, debido a que estas actividades no están incluidas en el registro de eventos y es necesario unir el inicio y fin de los flujos de actividades de alguna forma. Luego de realizar esta acción se generó un Modelo Fuzzy (Figura 3.5), donde se observan el flujo general del proceso. Este modelo arrojó un fitness de 0,9865, considerado elevado para la métrica.

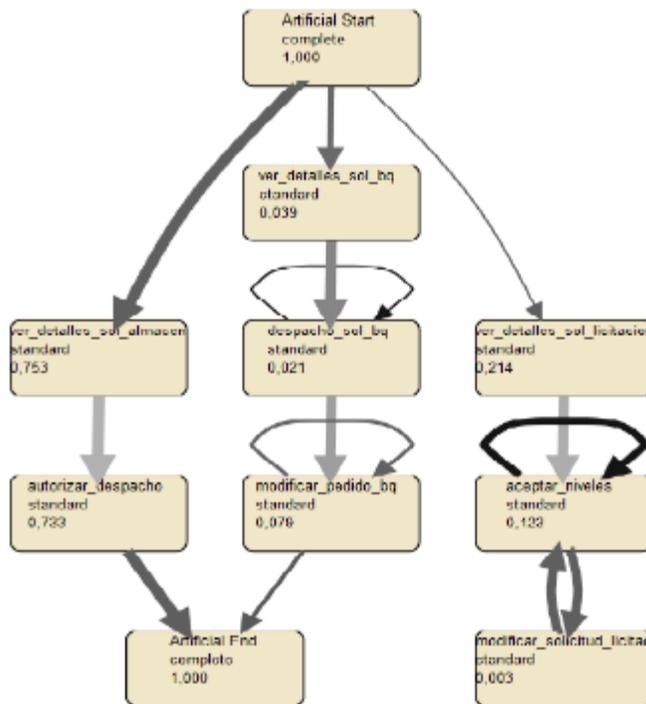


Figura 3.5. Modelo Fuzzy.

Fuente: Elaboración propia.

Una característica notable del modelo es la presencia de lazos en las actividades *despacho_sol_bq*, *modificar_pedido_bq* y *aceptar_niveles* (representados en azul en la Figura 3.6).

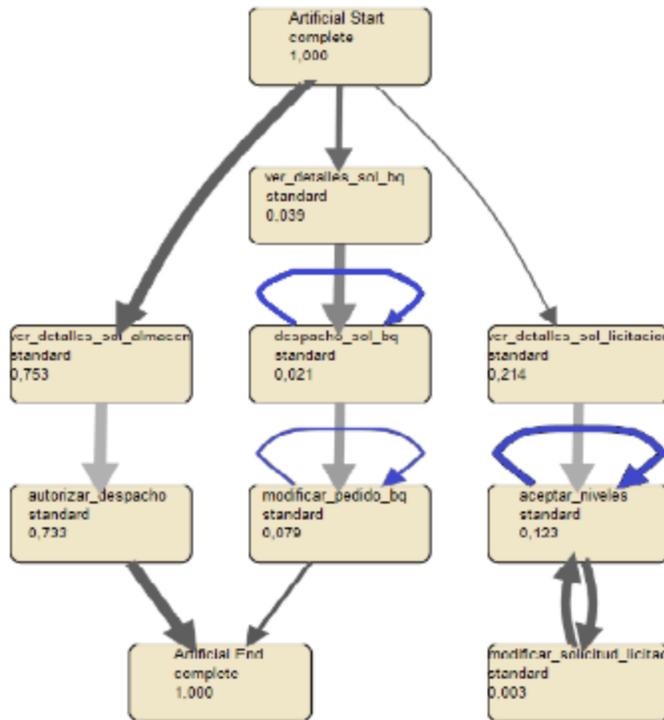


Figura 3.6. Modelo Fuzzy (Lazos en azul).

Fuente: Elaboración propia.

Otra de las posibilidades que ofrece este modelo es descubrir la frecuencia con que se ejecutan las actividades (representada por un número entre 0 y 1 dentro de las actividades) y la frecuencia de ocurrencia de los flujos (grosor de los arcos). En la Figura 3.7 se representa la frecuencia de aparición de las actividades y en la Figura 3.8 se resalta en rojo el flujo principal del proceso.



Figura 3.7. Frecuencia de ejecución de las actividades.

Fuente: Elaboración propia.

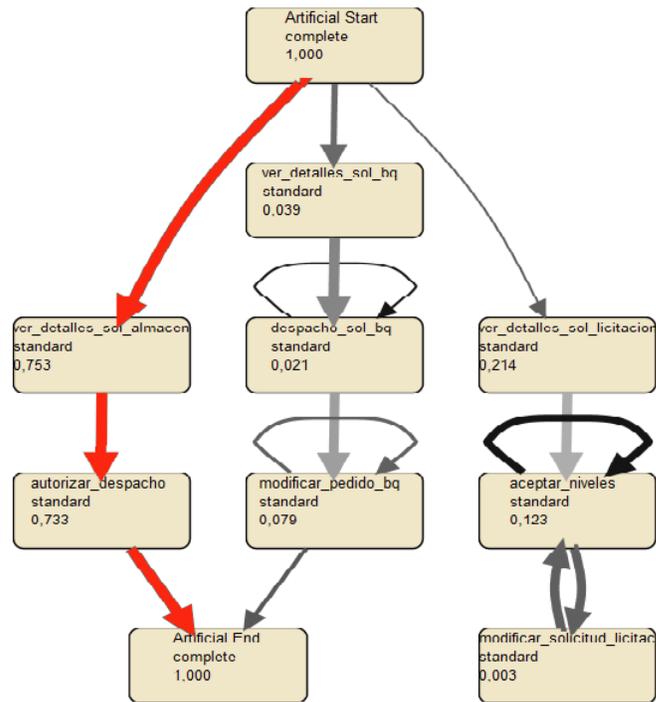


Figura 3.8. Modelo Fuzzy (Flujo principal en rojo).

Fuente: Elaboración propia.

Como se describió en el Capítulo 1, la técnica Alineación de Trazas (*Trace Alignment*), facilita determinar patrones de comportamiento en los datos. En la Figura 3.9 se observan los 9 flujos diferentes que se pueden dar en el proceso Solicitar producto y la codificación de cada actividad. Esta técnica arrojó un fitness de 98,22, el cual se considera elevado.

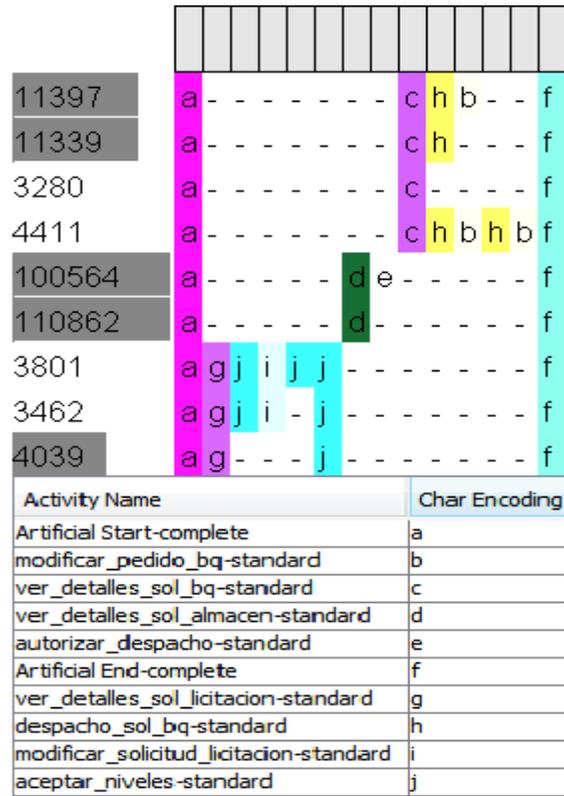


Figura 3.9. Modelo obtenido utilizando la técnica Alineación de Trazas (*Trace Alignment*).

Fuente: Elaboración propia.

Para descubrir el modelo primero se aplica la técnica "Mine a Process Tree using Inductive Miner", esta técnica permite generar un árbol de procesos que representa el flujo de actividades del proceso Solicitar producto del módulo Almacén. En la Figura 3.10 se representa el modelo BPMN obtenido con esta técnica. Una de las características que destaca de esta técnica es que ofrece la posibilidad de garantizar un fitness de valor 1, es decir que todas las trazas contenidas en el registro de eventos sean representadas en el modelo.

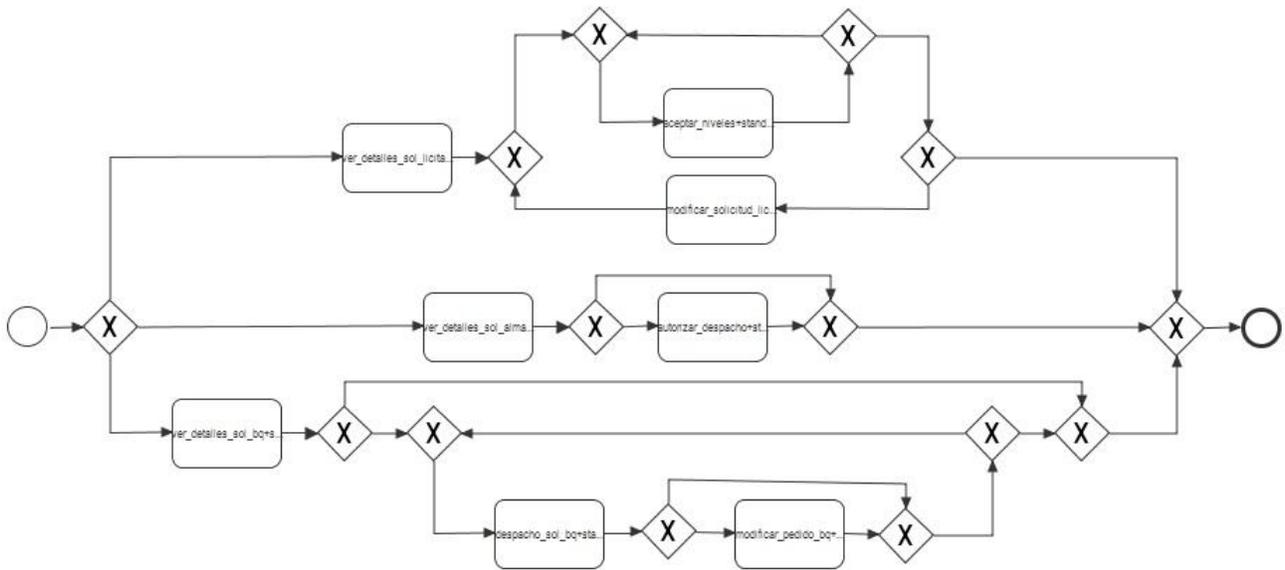


Figura 3.10. Modelo BPMN obtenido con la técnica "Mine a Process Tree using Inductive Miner"

Fuente: Elaboración propia.

Existe otro modelo que permite visualizar el flujo de actividades contenido en el registro de eventos denominado árbol de procesos. Este modelo como se muestra en la Figura 3.11 representa las actividades por niveles y como nodos en un árbol en el que se modela con una flecha apuntando a la derecha cuando existe una transición, si existe un lazo es representado con una x circundada por una flecha y una x sin flecha en caso de que sea una decisión.

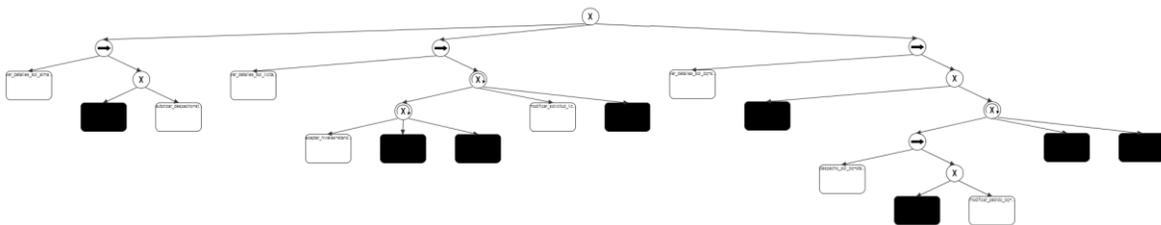


Figura 3.11. Árbol de proceso generado con la técnica "Mine a Process Tree using Inductive Miner"

Fuente: Elaboración propia.

Al comparar el Diagrama de Procesos del Negocio original del sistema representado en la Figura 3.1 y el modelo BPMN obtenido con la técnica "Mine a Process Tree using Inductive Miner" de la Figura 3.10, se comprueba si el flujo de actividades obtenido se corresponde con el diseñado originalmente. Tras

un minucioso análisis se comprobó que existen transiciones entre actividades del proceso Solicitar producto que nunca fueron ejecutadas por el sistema. Las transiciones no ejecutadas son las que van de la actividad *ver_detalle_sol_licitacion* y *modificar_sol_licitacion* a fin. En la Figura 3.12 se muestran estas actividades en el modelo obtenido encerradas en rectángulos rojos y en el Diagrama de Procesos del Negocios original se muestran las transiciones que nunca fueron ejecutadas por el sistema, marcadas en rojo igualmente.

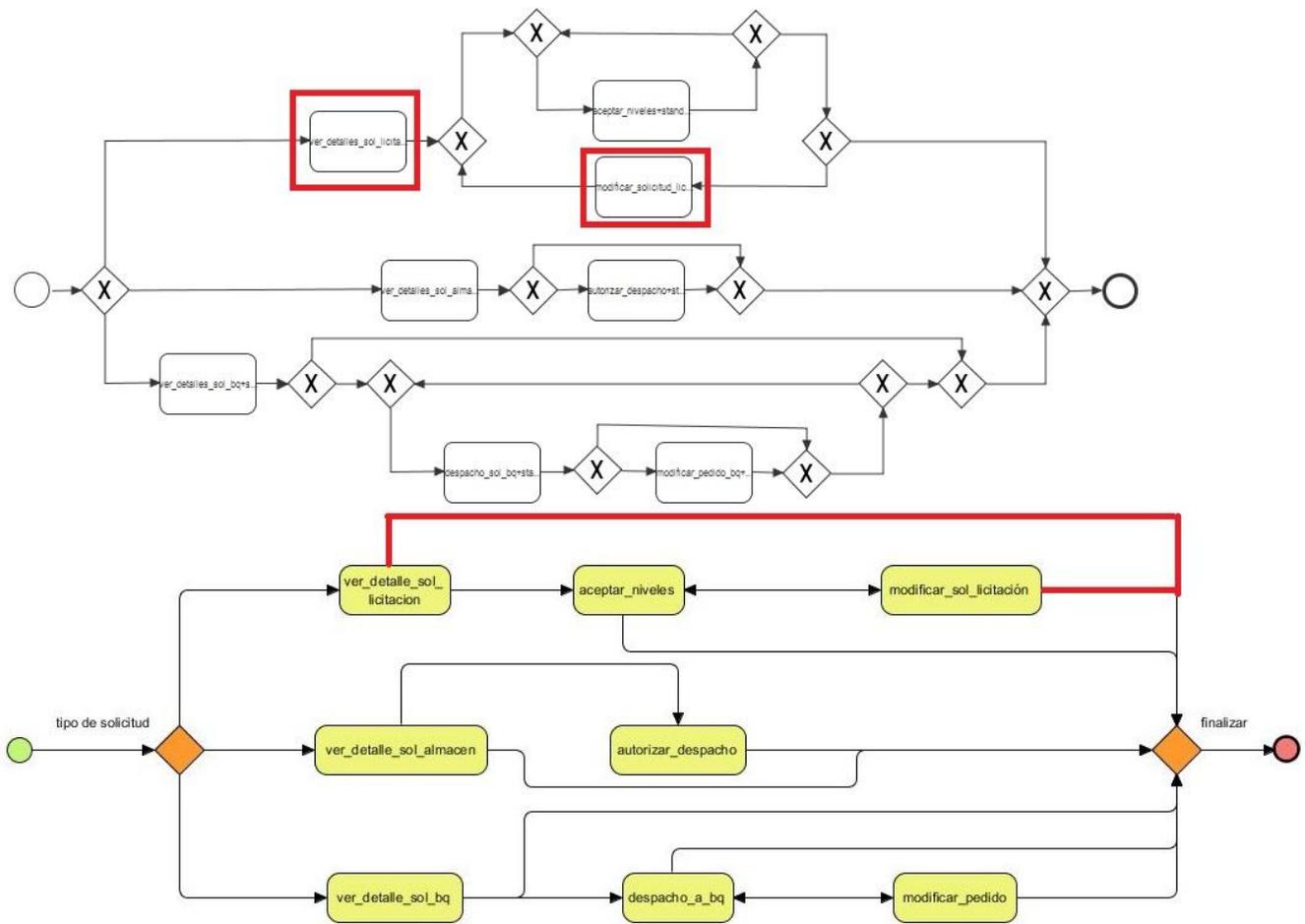


Figura 3.12. Comparación entre el modelo BPMN original y el generado por la técnica "Mine a Process Tree using Inductive Miner".

Fuente: Elaboración propia.

Posteriormente, se aplica la técnica "Convert Process Tree to Petri Net", generando una Red de Petri a la cual posteriormente se le integrará la perspectiva tiempo en la próxima etapa y que se muestra a continuación en la Figura 3.13. Se seleccionó esta técnica porque se puede garantizar un fitness de valor 1, en la Red de Petri obtenida porque el modelo BPMN del que se deriva cumple con esta característica también.

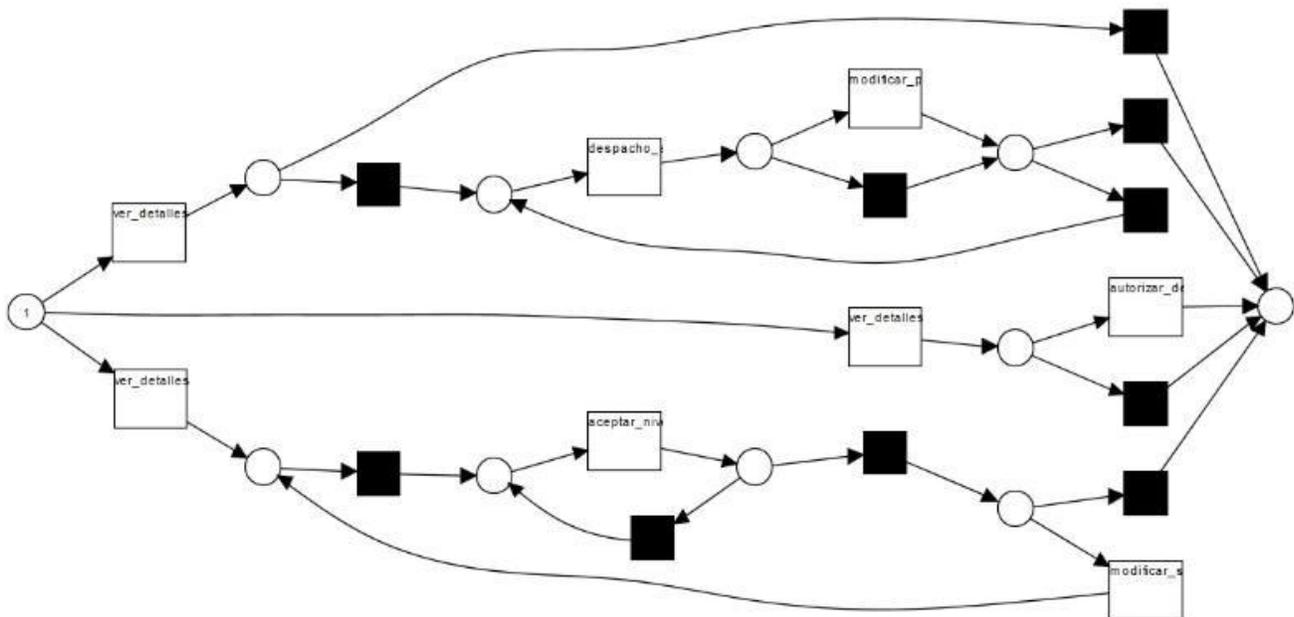


Figura 3.13. Aplicando la técnica "Convert a Process Tree to Petri Net"

Fuente: Elaboración propia.

3.2.4 Etapa 3 - Crear un modelo de proceso integrado

Fase 1. Selección de la técnica.

A partir de una lectura del libro *Discovery, Conformance and Enhancement of Business Process* (van der Aalst, 2011) se definió la utilización de la técnica "Replay a Registro on Petri Net for Performance/Conformance" basándose en las propiedades de su uso, planteadas por el autor del libro. Entre las características que definen esta técnica y que se explicarán en la fase 3, se encuentran:

- ❖ Detectar problemas.
- ❖ Extraer información de tiempo.

- ❖ Detectar cuellos de botella.
- ❖ Diagnosticar, predecir y recomendar.
- ❖ Generar modelo integrado mostrando los tiempos, frecuencias, etc.

Fase 2. Obtener el modelo integrado enfocado a la perspectiva tiempo.

En la Figura 3.14 se presenta el modelo integrado a la perspectiva tiempo obtenido al aplicar la técnica seleccionada en la fase anterior.

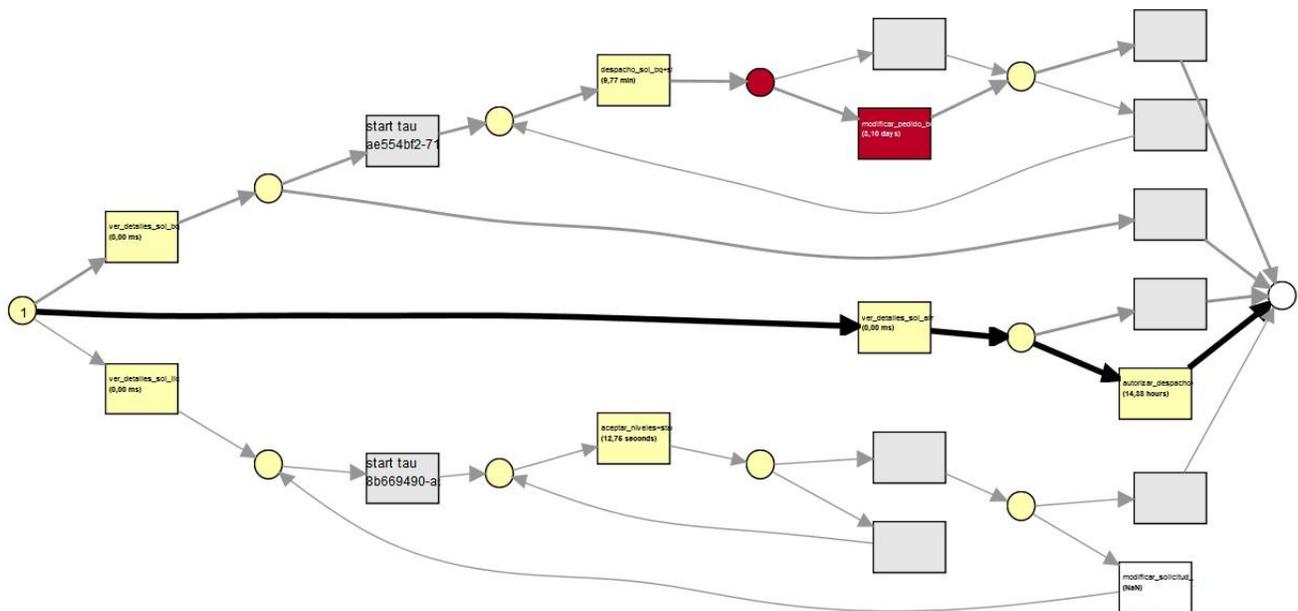


Figura 3.14. Modelo integrado a la perspectiva tiempo una vez aplicada la técnica "Replay a Registro on Petri net for Performance/Conformance"

Fuente: Elaboración propia.

Fase 3. Realizar el chequeo de conformidad al modelo integrado.

Para registrar una eficiente interpretación de los resultados se debe conocer primeramente cada uno de los atributos que conforman este modelo. Al analizar las transiciones, esta técnica establece que mientras más oscura y gruesa sea la transición, mayor cantidad de veces se han ejecutado las actividades que representan el flujo.

Esta representación permite en algunos casos definir a simple vista cuál es el flujo de actividades que más se ejecuta en el proceso del sistema bajo análisis. Como resultado de esta investigación se puede definir que en el proceso Solicitar producto el flujo de actividades más ejecutado es el compuesto por las actividades, *ver_detalle_sol_almacen* y *autorizar_despacho*.

Por otro lado, al analizar los nodos y revisar en la leyenda de técnica la escala de colores se puede identificar a simple vista que actividades son las más críticas y cuales funcionan correctamente. Se define como actividad crítica, aquella que sobrepasa el tiempo estimado de respuesta entre una actividad y otra. Este tiempo puede ser definido por el cliente o estimado por la técnica, teniendo en cuenta un estudio basado en el comportamiento de todas las ejecuciones anteriores de los flujos de actividades similares al que se encuentre bajo análisis.

La escala de colores utilizada por esta técnica es mostrada en la Figura 3.15, como se puede observar va desde un color amarillo claro hasta un color rojo vino. A mayor oscuridad en el color, mayor probabilidad de que exista un mal funcionamiento del sistema respecto al tiempo medio de estancia de las transiciones en una actividad antes de transportarse hacia otra (Adriansyah, 2012). Cuando se analiza el modelo integrado obtenido en la etapa anterior, se puede comprobar que existen seis nodos que presentan coloración blanca, uno amarilla clara y otro rojo vino. Los nodos de color blanco representan las transiciones que ocurrieron de forma eficiente y rápida. El nodo que aparece en amarillo claro, demuestra que existe una tardanza al recibir algunas de las respuestas del nodo que lo precede en el flujo de actividades correspondiente a él.

En este nodo queda representado el tiempo promedio que puede llegar a tardar un usuario o el sistema en *autorizar_despacho* una vez ha visto la actividad *ver_detalle_sol_almacen*. En este nodo se obtuvo que el tiempo mínimo de espera fue de 3.29 segundos, el tiempo máximo de espera fue de 5 meses y el promedio en tiempo de espera es de 14.38 horas. El último nodo a analizar *modificar_pedido_bq* presenta el promedio de tiempo de espera más crítico, siendo este de 3.15 días. Esto se debe a que su tiempo mínimo de espera es de 4.43 minutos y el tiempo máximo de 7.45 días, por lo que al calcular el promedio ese tiempo medio es el resultante.

Si se observa la diferencia entre los tiempos de espera máximo y mínimo entre las actividades *modificar_pedido_bq* y *autorizar_despacho* se nota que:

- ❖ El tiempo de espera mínimo de la primera actividad (3.29 segundos) es menor que el de la segunda (4.43 minutos).
- ❖ El tiempo de espera máximo de la primera actividad (5 meses) es mayor que el de la segunda actividad (7.45 días).
- ❖ El tiempo promedio de la primera actividad (14.38 horas) es menor que el de la segunda actividad (3,31 días).

Este último resultado se debe a la cantidad de veces que se ejecutan estas actividades en el registro de eventos. La actividad *autorizar_despacho* se ejecuta un total de 1128 veces mientras que la actividad *modificar_pedido_bq* solamente unas 15 veces. Cuando se calcula el tiempo promedio a estas actividades la cantidad de ejecuciones es un factor determinante en el resultado final. Un ejemplo de esto es cuando se tiene un caso que tiene pocas ejecuciones y un tiempo de espera máximo elevado respecto al tiempo mínimo como en la actividad *modificar_pedido_bq*.

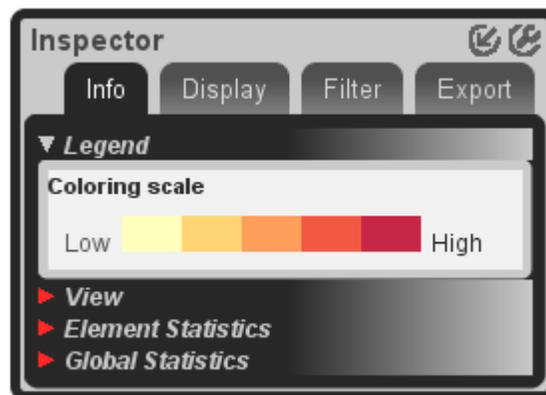


Figura 3.15. Escala de colores de la técnica "Replay a Registro on Petri net for Performance/Conformance"

Fuente: Elaboración propia.

3.2.5 Etapa 4 - Apoyo a las operaciones

En la Figura 3.16 se observa el resultado que arroja la técnica "Analyze using Dotted Chart"; en el eje superior se representan las marcas de tiempo. El tiempo de ejecución del proceso se extendió durante 17 meses, donde se observa tres momentos bien definidos, con comportamientos diferentes. La primera etapa de ejecución corresponde a la etapa de prueba, donde no se ejecutaron las actividades *ver_detalle_sol_almacen* y *autorizar_despacho*. Luego el sistema no registró ninguna ejecución del proceso por tres meses.

En la segunda y tercera etapa que se extendió durante un mes y medio y nueve meses y medio respectivamente con un tiempo de inactividad de 2 meses entre ellas, solamente se ejecutan las actividades *ver_detalle_sol_almacen* y *autorizar_despacho*. Por estas características este proceso presenta un comportamiento anómalo, ya que se dejan de ejecutar 6 actividades de las 8 que componen la versión final del proceso.

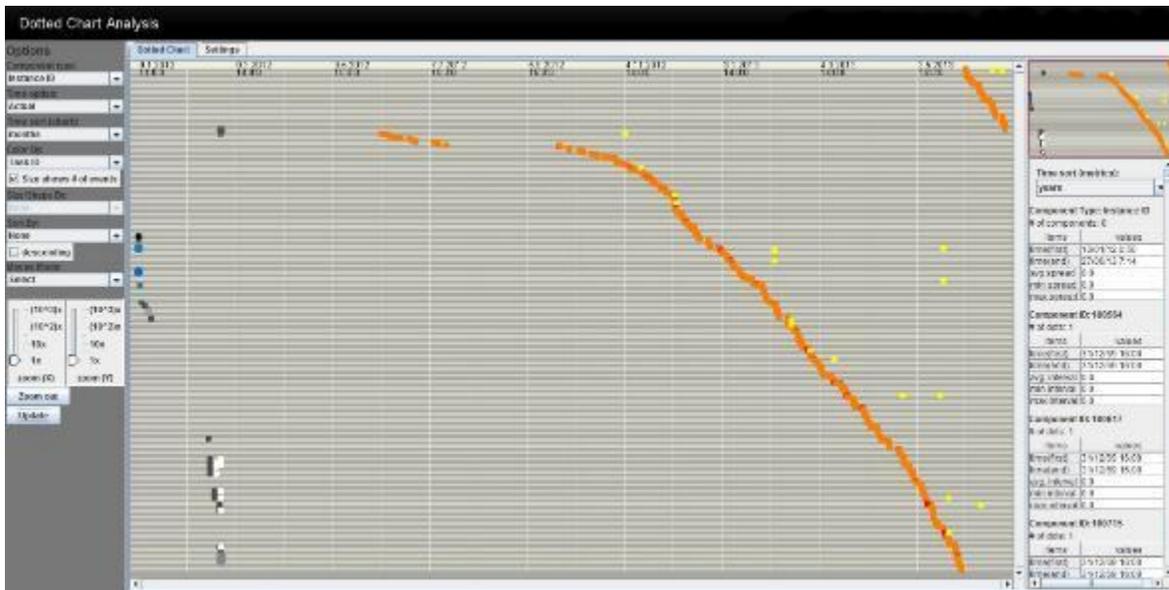


Figura 3.16. Resultado de la técnica "Analyze using Dotted Chart".

Fuente: Elaboración propia.

Como se observa en la Figura 3.17, otra de las características del proceso Solicitar producto es la presencia de ruidos que se resaltan con óvalos negros. Según (Yzquierdo, 2013), el ruido se refiere a situaciones donde el registro de eventos es incompleto, contiene errores o refleja un comportamiento excepcional. (van der Aalst, 2011) define al ruido como el 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. Este autor no hace referencia a los errores porque ningún registro de evento revela explícitamente los errores que contiene.

En el caso del modelo que se está analizando el ruido se refiere a comportamientos excepcionales más que errores. Esto se pone de manifiesto en los 3 períodos de tiempo donde el sistema deja de funcionar

y no se registran actividades. También, en la aparición de actividades aisladas que corresponden a atrasos en el cumplimiento de las solicitudes.



Figura 3.17. Formas de manifestación del ruido (circulado en negro) en el proceso Solicitar producto.

Fuente: Elaboración propia.

Al analizar las propiedades del flujo de actividades que más se ejecuta compuesto por la *ver_detalle_sol_almacen* y *autorizar_despacho*, utilizando el módulo inspector (Figura 3.18) se observan desviaciones en el tiempo de respuesta de las solicitudes. Según este módulo 1128 casos se ajustaron perfectamente y 84 no se ajustaron. 1176 casos iniciaron correctamente, con un tiempo promedio de cumplimiento de 14,31 horas. El tiempo mínimo para autorizar un despacho fue de 0,00 milisegundos y el tiempo máximo de 5 meses. La desviación estándar, que representa la dispersión de las actividades con respecto a la media, es de 7,26 días.

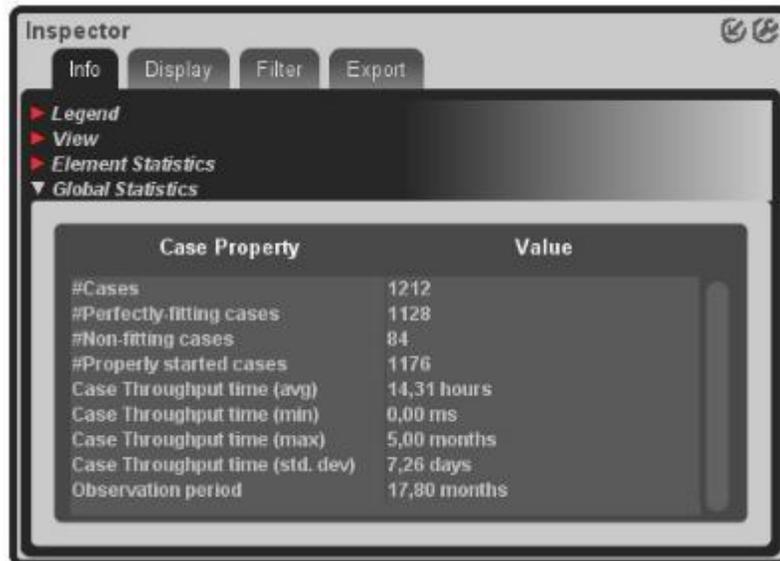


Figura 3.18. Propiedades del flujo de actividades *ver_detalle_sol_almacen* y *autorizar_despacho*.

Fuente: Elaboración propia.

En la Figura 3.19, se ejemplifica la desviación temporal a la hora de autorizar un despacho. Los puntos rojos representan la actividad *ver_detalle_sol_almacen*, los amarillos *autorizar_despacho* y los anaranjados la correcta secuencia de ejecución de ambos.

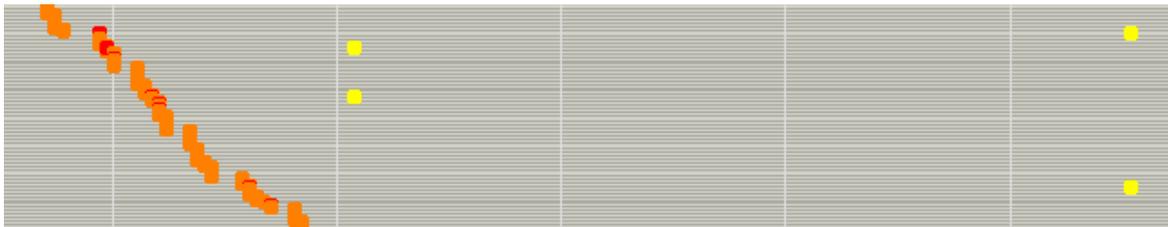


Figura 3.19. Desviación temporal a la hora de autorizar un despacho.

Fuente: Elaboración propia.

Conclusiones del capítulo

1. El modelo obtenido describe un proceso estructurado o Lasaña debido a las pocas actividades y relaciones que se establecen.
2. Se detectaron desviaciones en la conformidad del modelo, causado principalmente por la presencia de ruido en el proceso.
3. Se analizaron 1212 casos, de los cuales 84 no se ajustaban correctamente al funcionamiento estándar del sistema.

CONCLUSIONES

Se realizó un análisis del estado del arte, sobre las diferentes aplicaciones de la minería de procesos. A partir de este análisis se propusieron objetivos que fueron cumplidos y se arriba a las siguientes conclusiones:

- La selección de las técnicas estuvo determinada por las características del registro de eventos y de la información que se puede extraer de este.
- El resultado de la aplicación de este modelo quedó expresado en los modelos de descubrimiento y conformidad obtenidos.
- Estos modelos arrojaron resultados como la detección de retrasos en el cumplimiento de diferentes instancias de la actividad *autorizar_despacho*. Además el proceso presenta desviaciones en el patrón de comportamiento en los 17 meses de ejecución del mismo.
- Las transiciones que van desde la actividad *ver_detalles_sol_licitacion* y *modificar_solicitud_licitacion* al final del proceso nunca son ejecutadas.
- La actividad *modificar_pedido_bq* representa un cuello de botella en el funcionamiento del sistema.

RECOMENDACIONES

1. Reestructurar el registro de trazas del sistema siguiendo un enfoque basado en procesos.
2. Continuar con la aplicación de las técnicas de mejora al Sistema de Información Hospitalaria.

REFERENCIA BIBLIOGRÁFICA

Adriansyah, A. & Buijs, J.C.A.M. (2012). Mining Process Performance from Event Registros The BPI Challenge 2012 Case Study. Eindhoven University of Technoregistry, Netherlands.

Afyouni, I., Ray, C. & Ilarri, S. (2013). A PostgreSQL extension for continuous path and range queries in indoor mobile environments. United States.

Barros, O. (2012). Ingeniería de Negocios: Diseño Integrado de Negocios, Procesos y aplicaciones TI.

Bose, R.P. & van der Aalst, W.M.P. (2011). Process Diagnostics Using Trace Alignment: Opportunities, Issues, and Challenges. Eindhoven University of Technoregistry, Netherlands.

Bozkaya, M. (2009). Process Diagnostics: a Method Based on Process Mining.

Bozkaya, M., Gabriels, J. & van der Werf, J. M. (2009). Process Diagnostics: a Method Based on Process Mining. Trabajo presentado en The International Conference on Information, Process, and Knowledge Management, Cancun, México.

Bratosin, C. (2011). Grid Architecture for Distributed Process Mining. Tesis Doctoral, Eindhoven University of Technoregistry, Netherlands.

Castellanos, M., Medeiros, A.K.A., Mendling, J., Weber, B. & Weijters, A.J.M.M. (2009). Business Process Intelligence. University of Innsbruck, Australia.

Chamorro, M.C. (2013). Método para aplicar minería de procesos a la distribución de bebestibles no alcohólicos. Santiago de Chile, Chile.

Gao, X. (2013). Towards the Next Generation Intelligent BPM. In the Era of Big Data. Department of Management Information System. China Mobile Communications Corporation, Beijing 100033, China.

Harmen, R. (2012). A practitioner's guide for process mining on ERP Systems- The case of sap order to cash. Eindhoven University of Technoregistry, Netherlands.

Jans, M. (2011). Process Mining of Event Registros in Internal Auditing: A Case Study. Trabajo presentado en The 2nd International Symposium on Accounting Information Systems, Italy.

Laszewski, T. & Nauduri, P. (2012). Oracle Client/Server Modernization. Waltham, United States.

-
- Leoni, M. & van der Aalst, W.M.P. (2012).** Aligning Event Registros and Process Models for Multi-Perspective Conformance Checking: An Approach Based on Integer Linear Programming. Eindhoven University of Technoregistry, Netherlands.
- Machío, V. (2010).** Sociedad de la información en España: competencia autonómica de orientación europea. España.
- Mans, R.S., Schonenberg, M.H., Song, M., van der Aalst, W.M.P. & Bakker, P.J.M. (2008).** Process mining in healthcare. A Case Study. Eindhoven University of Technoregistry, Netherlands.
- Medeiros, A.K.A. (2006).** Genetic Process Mining. Eindhoven University of Technoregistry.
- Muñoz-Gama, J. (2010).** ETConformance. Universidad Politécnica de Cataluña, España.
- NC/ISO. (2005).** Sistemas de Gestión de la Calidad-Fundamentos y Vocabularios, NC ISO 9000: 2005.
- Pérez, D., Yzquierdo, R., Silveiro, R. & Lazo, M. (2012).** Utilización de técnicas de minería de proceso en el entorno empresarial cubano. Universidad de las Ciencias Informáticas, La Habana, Cuba.
- Recker, J., Rosemann, M., Indulska, M. & Green, P. (2007).** Business Process Modeling: A Maturing Discipline?
- Rozinat, A. & van der Aalst, W.M.P. (2006).** Conformance testing: Measuring the fit and appropriateness of event registros and process models. Business Process Management Workshops.
- Rozinat, A. & van der Aalst, W.M.P. (2008).** Conformance Checking of Processes Based on Monitoring Real Behavior.
- Rozinat, A., Mans, R. S., Song, M. & van der Aalst, W.M.P. (2009).** Discovering Simulation Models. Information Systems.
- Saravanan, M.S. & Rama, R.J. (2010).** Process mining in healthcare using control flow perspective: a case study. India.
- Seguel, R. (2008).** Process Mining. Eindhoven University of Technoregistry, Netherlands.
- Song, M. & van der Aalst, W.M.P. (2007).** Supporting Process Mining by Showing Events at a Glance. Trabajo presentado en la 17th Annual Workshop on Information Technoregistroies and Systems, Montreal, Canadá.

-
- van der Aalst, W. M. P. & Weijters, A. (2004).** Process mining: a research agenda. Computers in Industry.
- van der Aalst, W.M.P. (2011).** Process Mining. Discovery, Conformance and Enhancement of Business Processes. London New York: Springer.
- van der Aalst, W.M.P., Adriansyah, A., Medeiros, A.K.A. et al. (2011).** Process Mining Manifesto. IEEE.
- van der Aalst, W.M.P., et al. (2009).** "ProM: The Process Mining Toolkit," in Proceedings of BPM (Demos)'2009, Ulm, Germany. CEUR-WS.org, vol. 489.
- van der Aalst, W.M.P., Reijers, H.A., Weijters, A.J.M.M., van Dongen, B.F., Alves de Medeiros, A.K., Song, M. & Verbeek, H.M.W. (2007).** Business process mining: An industrial application. Eindhoven University of Technoregistry, Netherlands.
- van der Aalst, W.M.P., ter Hofstede, Arthur, H. M. & Weske, M. (2003).** Business Process Management: International Conference. Eindhoven University of Technoregistry, Netherlands.
- van Dongen, B. & van der Aalst, W. M. P. (2005).** A meta model for process mining data. Proceedings of the CAiSE.
- van Dongen, B. (2007).** Process Mining and Verification, PhD Thesis. Eindhoven University of Technoregistry, Netherlands.
- van Doreman, B. (2012).** Process Mining in Healthcare Systems: An Evaluation and Refinement of a Methodoregistry. Eindhoven University of Technoregistry, Netherlands.
- van Giessel, M. (2004).** Process mining in SAP R/3. Tesis de Maestría, Eindhoven University of Technoregistry, The Netherlands.
- Verbeek, H.M.W. (2010).** ProM 6 Tutorial.
- Verbeek, H.M.W. (2011).** XES, XESame, and ProM 6.
- Yanga, W. & Hwang, S. (2006).** A process-mining framework for the detection of healthcare fraud and abuse. Taiwan.
- Yaofei, C., Huantong, C. & Yinghua, N. (2011).** Application of ADOX in SQL Server Skill Assessment. Jinhua Polytechnic, Jinhua, China

Yzquierdo, R. (2012). Modelo para la estimación de información ausente en las trazas usadas en la minería de proceso. Tesis de Doctorado, Universidad de las Ciencias Informáticas, Cuba.

WEBGRAFIA

Disco. (n.d.). Extraído el 24 de abril de 2014 desde <http://www.fluxicon.org/>

EMS. (n.d.). Extraído el 1 de mayo de 2014 desde <http://www.sqlmanager.net/>

Microsoft. (n.d.). Extraído el 28 de abril de 2014 desde <http://www.microsoft.com/>

Nitro. (n.d.). Extraído el 24 de abril de 2014 desde <http://www.fluxicon.org/>

Oracle. (n.d.). Extraído el 28 de abril de 2014 desde <http://www.oracle.com/>

PgAdmin. (n.d.). Extraído el 1 de mayo de 2014 desde <http://www.pgadmin.org/>

PostgreSQL. (n.d.). Extraído el 28 de abril de 2014 desde <http://www.postgresql.org/>

ProM. (n.d.). Extraído el 25 de abril de 2014 desde <http://www.promtool.org/prom6>

XESame. (n.d.). Extraído el 20 de abril de 2014 desde <http://www.processmining.org/>

GLOSARIO DE TÉRMINOS

BPMN: notación gráfica estandarizada que permite el modelado de procesos de negocio, en un formato de flujo de trabajo.

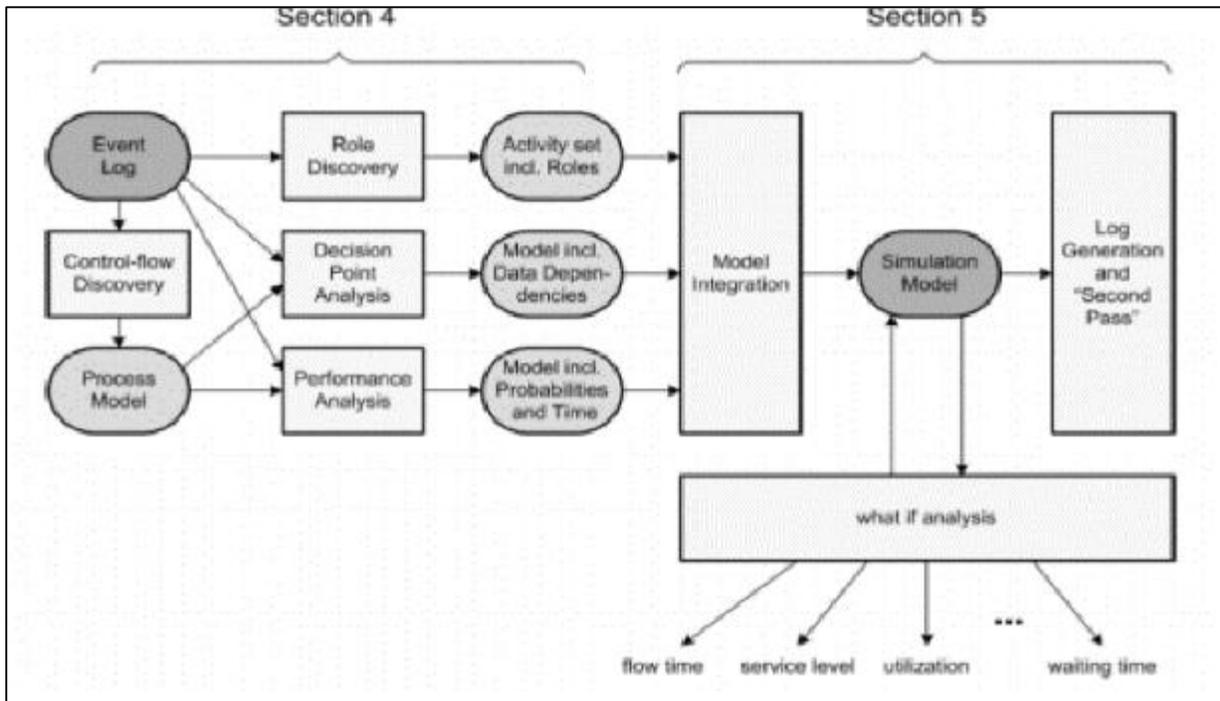
Chequeo de conformidad: uno de los tres tipos de minería de procesos junto con el descubrimiento y la mejora; permite identificar desviaciones entre el funcionamiento real del sistema y el modelado originalmente.

Minería de Procesos: disciplina de investigación que se encarga de extraer información útil y oculta de los registros de trazas generados por los Sistemas de Gestión de Información.

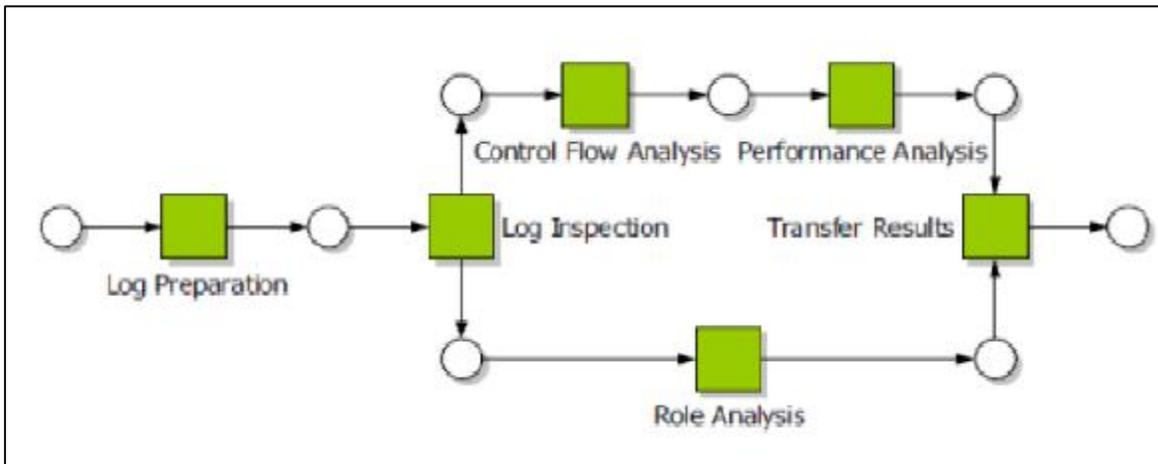
Registro de eventos: Mecanismo utilizado por los Sistemas de Gestión de Información para almacenar la ejecución de los procesos en el sistema; se registran en forma de trazas

XES: Estándar implantado en el 2011 para almacenar los registros de eventos y servir como entrada a las herramientas utilizadas en la minería de procesos.

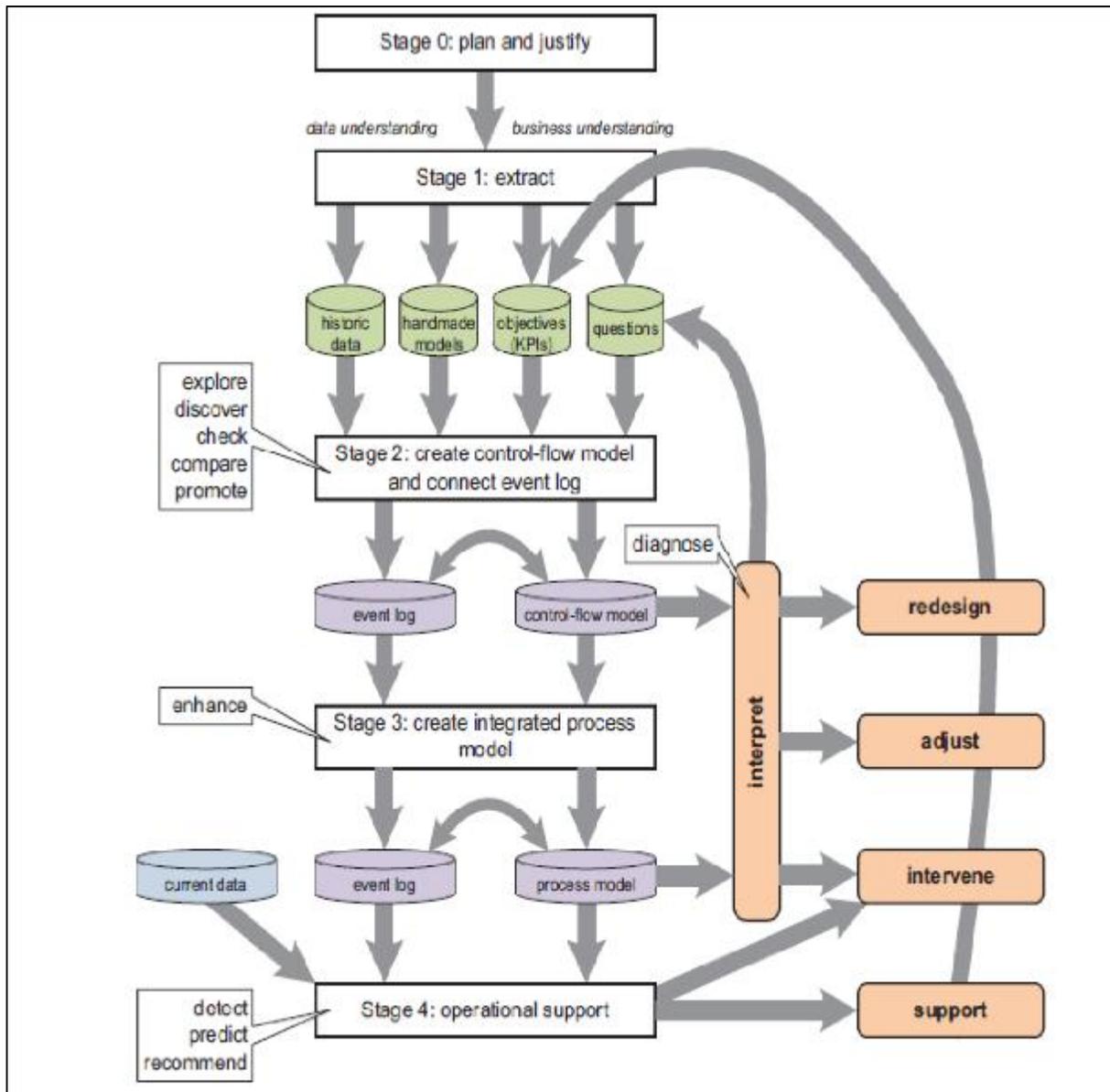
ANEXOS



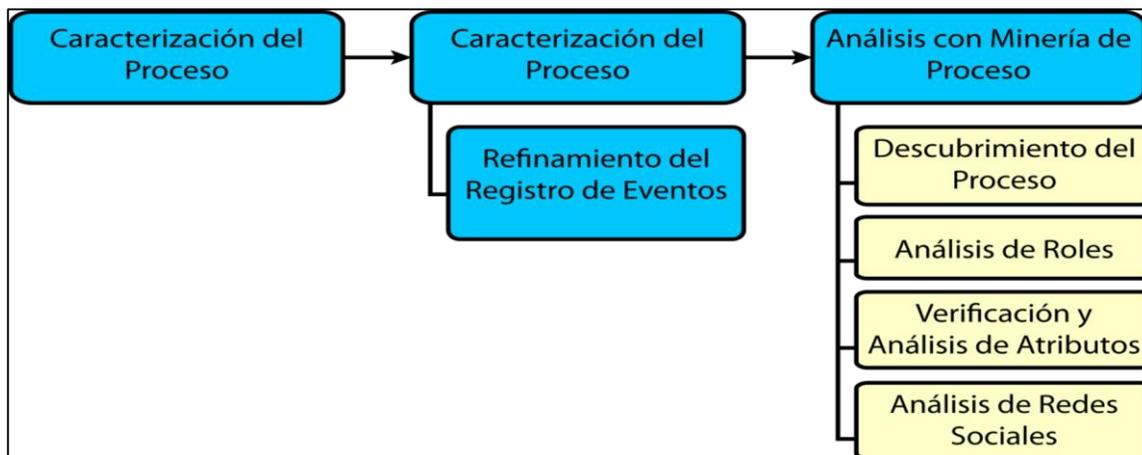
Anexo 1. Modelo de Rozinat



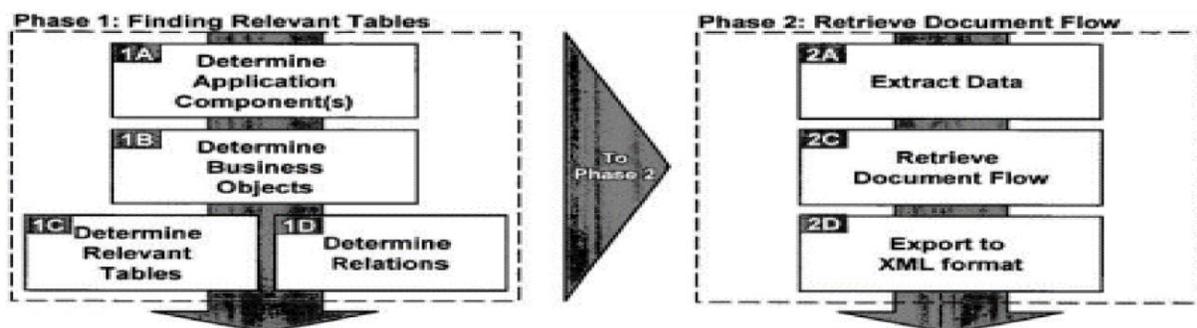
Anexo 2. Modelo de Bozkaya



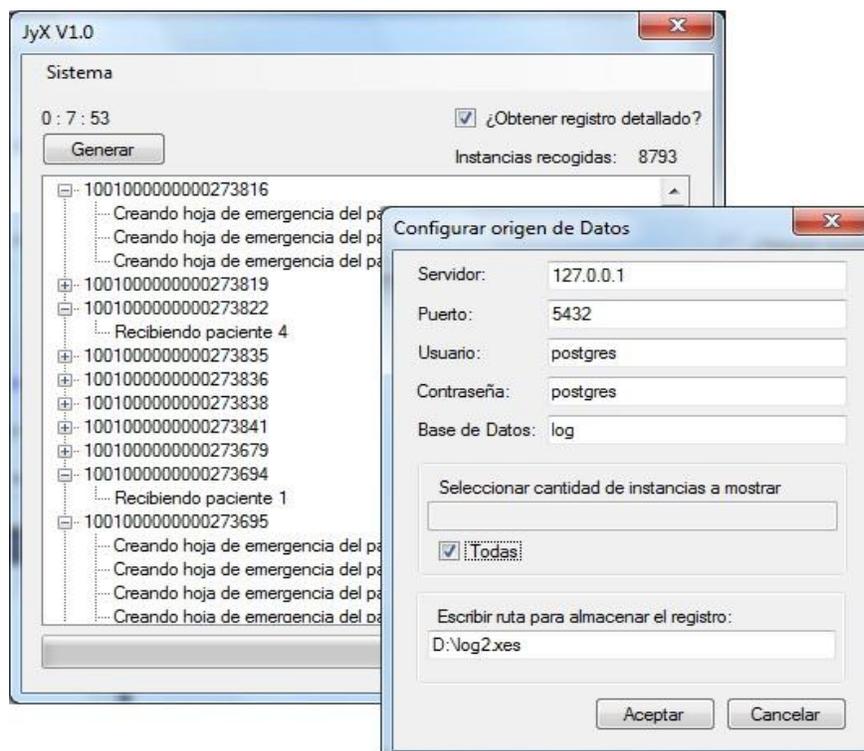
Anexo 3. Modelo de van der Aalst



Anexo 4. Modelo de Jans.



Anexo 5. Modelo de van Giessel



Anexo 6. Interfaz para Windows de la herramienta JyX v1.0