

# Análisis de frecuencia de ejecución de procesos hospitalarios aplicando la técnica Fuzzy Miner de minería de procesos

1

2

*Abstract— Process mining allows discover the real execution from systems. This technology can analyze event logs from the application of existing techniques in developed in plugins. Event logs are the product of extract and transform execution traces of systems, and are ideal to gain knowledge the performance of the institution. One of the plugins that have been developed to obtain process models is Fuzzy Miner. This technique is useful to identify behaviors like frequency and deviations in the process flow. An interpretation of the presence low frequency is the noise; it is occurring as anomalous activities in the process that undermine the proper functioning of the system. The research focused on customizing Fuzzy Miner technique to identify frequency behaviors in the process of Hospital Information System of the Center for Medical Informatics. As a result, it is expected that the implementation of the proposed solution to identify the frequency in event logs generating models that allow a better understanding and usability for non-experts in Process Mining and support decision making in hospitals institutions.*

*Keywords— fuzzy miner, frequency identification, process mining, process model, technique.*

## I. INTRODUCCIÓN

Con el fin de obtener resultados eficientes, constituye una prioridad mejorar la administración y desempeño organizacional a través de la incorporación de la gestión o enfoque basado en procesos. El enfoque basado en procesos es un principio de gestión básico y fundamental para la obtención de resultados. Este principio sostiene que “un resultado se alcanza más eficientemente cuando las actividades y los recursos se gestionan como un proceso” [1]. Entiéndase por proceso la actividad o conjunto de actividades que utiliza recursos, y que se gestiona con el fin de permitir que los elementos de entrada se transformen en resultados [2].

En este sentido, las instituciones del sector de la salud, se han ido modernizando con los sistemas y las tecnologías de información que han surgido [3] [4], permitiéndoles avanzar en la prestación de servicios, sin embargo esto va generando

**Digital Object Identifier:** (to be inserted by LACCEI).

considerables volúmenes de datos derivados de las

actividades de los procesos clínicos que necesitan ser analizados para que su aporte sea verdaderamente significativo. Como alternativa para el análisis de procesos surge la minería de procesos [5].

La minería de procesos posibilita entender cómo son ejecutados en realidad los procesos en sistemas automatizados [6], pretende descubrir, monitorear y mejorar procesos reales mediante la extracción de conocimiento de los registros de eventos disponibles en los sistemas información [7]. Los *logs*<sup>1</sup> son utilizados para realizar los tres tipos de Minería de Procesos existentes son: descubrimiento, conformidad y mejora, siendo el descubrimiento de procesos el más investigado.

Entre los autores más destacados en investigaciones acerca de minería de procesos en el área de salud, se destacan van der Aalst [8] y Mans [9] [10], quienes destacan los beneficios de su aplicación a este sector de la sociedad. En investigaciones realizadas, se evidencia que es una tecnología efectiva para el análisis de procesos hospitalarios. Su uso en casos reales de procesos de ginecología y oncología [10] permitió optimizar, a partir de un gráfico de puntos, la trayectoria de los pacientes por el proceso de atención. Tras descubrir horarios de mayor afluencia en el área de Emergencias [9], permitió controlar y destinar recursos a este servicio.

Una de las técnicas de minería de procesos es Fuzzy Miner [11], la cual es capaz de detectar y manejar con éxito la presencia de ruido en los registros de eventos. Para esta investigación se asume la definición de ruido: “Comportamiento reflejado en las trazas y que rara vez ocurre, que es excepcional o poco frecuente, es decir, que no se corresponde con el comportamiento típico observado en el proceso” propuesta por van der Aalst [8].

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

---

<sup>1</sup> logs: nace del término en inglés log, que significa registro.

cual agiliza la obtención de los resultados y proporciona una gran versatilidad a la hora de realizar análisis y toma de decisiones. Es una alternativa útil para realizar análisis de frecuencia de la ejecución de los procesos, así mismo es posible identificar las desviaciones existentes en el flujo de actividades.

En el Centro de Informática Médica (CESIM) de la Universidad de las Ciencias Informáticas (UCI) se desarrolla el Sistema de Información Hospitalaria XAVIA HIS (HIS, por sus siglas en inglés). Sobre el sistema se desarrolló un componente para la extracción y transformación de las trazas almacenadas en registros de eventos [12]. Sin embargo para su análisis es necesario hacer uso de herramientas externas al dominio hospitalario como ProM [13]. ProM es una herramienta académica de libre acceso que soporta más de 600 técnicas con capacidades de minería de procesos.

En investigaciones realizadas [14] se ha comprobado que los modelos generados por la herramienta ProM dificultan la interpretación y el entendimiento de los usuarios no expertos en el área de minería de procesos. Los análisis con la técnica Fuzzy Miner sobre comportamientos que son infrecuentes o variables puede llevar a resultados erróneos. Además, su uso es complejo debido a que cuenta con varias métricas y parámetros de necesaria configuración y que no son del dominio del personal sanitario.

Los elementos de variabilidad como la frecuencia de ejecución y las desviaciones tienen un impacto estructural considerable sobre los modelos que se obtienen, debido a que puede influir en las decisiones que se toman sobre determinados servicios. La presencia de desviaciones en el modelo constituye un indicador de posibles fraudes o violaciones en las políticas lo cual puede ser útil su identificación por las organizaciones.

La presente investigación se plantea como objetivo realizar análisis de frecuencia y desviaciones en la ejecución de los procesos del XAVIA HIS a partir de la integración de la técnica Fuzzy Miner al sistema.

## II. FUZZY MINER

La técnica Fuzzy Miner está basada en el algoritmo Fuzzy Miner creado por Günther cofundador de Fluxicon en 2007. Está basada en medidas de significación<sup>2</sup> y correlación<sup>3</sup> para visualizar el comportamiento en los registros de eventos. Es el primer algoritmo que se ocupa directamente de los problemas de números grandes de actividades y comportamiento altamente no estructurado.

Según Günther [11] está caracterizada por la presencia de dos tipos de nodos; nodos primitivos que se refieren a una

<sup>2</sup>significación: importancia de cada una de las actividades con respecto a los demás.

<sup>3</sup>correlación: es medida por el tiempo de ejecución de dos actividades.

tarea y nodos que se refieren a un conjunto de tareas o clúster (*cluster*, por su nombre en inglés), considerando que cada tarea pertenece a un único nodo y es empleada en el diagnóstico del registro de eventos donde posibilita realizar análisis preliminares al descubrimiento.

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

La abstracción consiste en omitir de la visualización la información que en determinado contexto es insignificante. En determinados componentes se destacan características como color, contraste, saturación y tamaño para facilitar la visualización de la información relevante. Por su parte la personalización permite configurar la información que se muestra de acuerdo a determinados criterios [15].

Es una técnica de descubrimiento y este algoritmo es bastante aceptado como técnica de exploración de datos. Es aconsejable utilizar esta técnica cuando tenga los datos de registros complejos y no estructurados o cuando se quiere simplificar el modelo de una manera interactiva. Con su aplicación se obtiene un modelo basado en gráficos capaz de proporcionar una vista de alto nivel de un proceso, con la abstracción de los detalles no deseados. Fuzzy Miner [15] haciendo uso de umbrales de frecuencia configurables por el usuario descarta del modelo el ruido.

### A. Métricas y configuraciones

La técnica Fuzzy Miner fue desarrollada con el fin de optimizar y mejorar los procesos. Hace uso de técnicas para la extracción y agrupación de actividades con el objetivo de representar un proceso que sea comprensible por los analistas de procesos. Fuzzy Miner cuenta con una serie de métricas configurables derivadas de las mediciones del mismo para generar los modelos. A continuación se resumen las configuraciones presentadas en tres artículos en particular [15] [16] [17].

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

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

- Significado de frecuencia: Esta dado por la cantidad de veces que se repite una actividad con respecto a todas las demás en un registro de eventos. La métrica es normalizada, por tanto, la actividad que más se repite toma el valor de uno mientras que el valor de frecuencia de las demás es calculado a partir de ella.

- Significado de enrutamiento: Esta dado por el balance que exista entre los arcos que entran a un nodo y los que salen de él. Mientras mayor sean las conexiones con otras actividades (mayor cantidad de arcos que entran y salen) será mayor significado de enrutamiento.

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

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

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

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

- La correlación de proximidad: Evalúa la ocurrencia de un evento con respecto a otro, mientras menor tiempo de ejecución mayor correlación. Esta métrica es importante para identificar cluster de eventos que corresponde a una sola actividad lógica.
- La correlación originadora: Se tiene en cuenta el nombre de la persona que llevo a cabo la sub-secuencia de dos eventos. Mientras mayor similitud tengan los nombres mayor será el valor de correlación.
- La correlación de punto final: Es similar a la correlación originadora sin embargo la comparación que se tiene en cuenta es el nombre del recurso con el nombre de la actividad entre las dos sub-secuencia de actividades. Mientras más similares sean los nombres mayor será el valor de correlación.
- La correlación tipo de datos: En la mayoría de los registros, se incluyen atributos adicionales. Estos atributos son interpretados en el contexto en el que son usados. Esta métrica evalúa cada actividad. Mientras mayor cantidad de atributos existan y sean comunes entre las dos actividades, mayor será su correlación.
- La correlación valor de datos: Es semejante a la correlación tipo de datos pero no tiene en cuenta el tipo de datos si no los valores que contengan los atributos que son comunes. Un pequeño cambio en un atributo comprometerá el valor de la correlación.

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

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

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

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

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

- Configuración del punto de medida

En el área de configuración del punto de medida cuenta con un histograma que visualiza el número de puntos de medición por eventos (por ejemplo, el número de barras de histograma) y su factor de evaluación (altura de las barras). La Maximal event distance (distancia máxima de eventos) define el número de puntos de medida.

- Atenuación

La atenuación más simple es la Linear attenuation (atenuación lineal), que garantizará la atenuación con la distancia de eventos. La atenuación  $N^{\text{th}}$  root whit radical (raíz  $n$ -ésima con radicales) permite la atenuación negativa exponencial para configurar la función  $N^{\text{th}}$  root. Un valor relativamente alto atenuará progresivamente los puntos de mayor distancia de medida, que es útil cuando desea centrarse en las relaciones a corto plazo. Un valor relativamente bajo ayuda a identificar mejor las actividades que constituyen ruido en el registro de eventos [16].

## B. Ventajas y limitaciones

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

Entre sus principales limitaciones se encuentran las iteraciones, las múltiples instancias y el comportamiento dentro de un cluster por lo que se requiere de un estudio más detallado. La comparación con otros algoritmos también está limitada en el sentido de que un modelo difuso no puede

transformarse a una Red de Petri [18] y por tanto no permite una evaluación comparativa con otras técnicas de descubrimiento de procesos. El modelo difuso que genera la técnica no se puede convertir a otros tipos de lenguajes de modelado de proceso pero sí se puede animar para conseguir una sensación dinámica del proceso.

### III. PROCEDIMIENTO PARA LA INTEGRACIÓN DE LA TÉCNICA AL XAVIA HIS

La Figura 1 representa el modelo conceptual de la integración y funcionamiento de la técnica Fuzzy Miner en el XAVIA HIS. El objetivo es integrar al HIS una personalización de la técnica Fuzzy para visualizar los modelos de proceso y favorecer los análisis de frecuencia y desviaciones. Para ello utilizará los registros de eventos generados por el componente que tiene el sistema para la extracción y transformación de las trazas. La integración contribuirá a mejorar el uso de la técnica para los no expertos incorporando un resumen de la ejecución del proceso en lenguaje natural. Las configuraciones técnicas serán invisibles al usuario, la interfaz de la técnica solo define parámetros comprensibles que sugieran tipos de análisis apropiados.

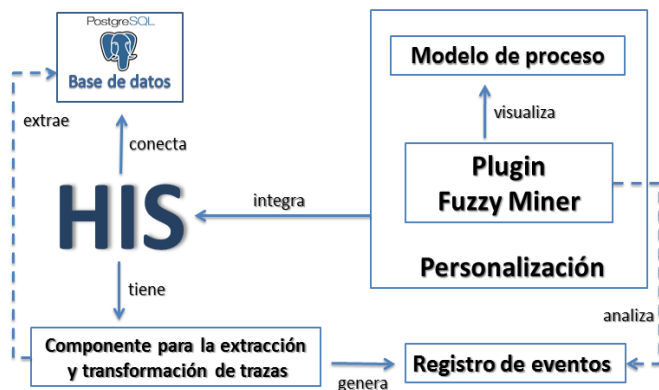


Fig. 1 Modelo conceptual de la integración de la técnica Fuzzy Miner al XAVIA HIS. Fuente: Elaboración propia.

El procedimiento para la integración de Fuzzy Miner de tiene como objetivo proveer al proceso de desarrollo de software del HIS de un *plugin* que contribuye a la visualización de un modelo de proceso entendible con la abstracción de los detalles no deseados. Para su aplicación deben cumplirse las premisas establecidas para el componente de extracción de registros de eventos del HIS: el log de eventos debe estar en formato \*.XES<sup>4</sup> y registrar como

<sup>4</sup>.XES: es un estándar XML para los registros de eventos. El estándar ha sido adoptado por la IEEE Task Force on Process Mining como el formato de intercambio de registros de eventos por defecto.

información mínima asociada al evento un identificador, el originador y la marca de tiempo.

El procedimiento propone tres fases para su aplicación: “Determinar las configuraciones necesarias para el HIS”, “Integrar la técnica Fuzzy Miner al HIS” y “Mostrar el modelo desde el HIS”.

#### Fase 1: Determinar las configuraciones necesarias para el HIS

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

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

```

<trace>
  1 <string key="concept:name" value="17438"/>
  2 <string key="description" value="Simulated process instance"/>
<event>
  3 <string key="org:resource" value="mirandae@pdvsa.com"/>
  4 <date key="time:timestamp" value="2012-06-20T15:47:03.703.000+01:00"/>
  5 <string key="concept:name" value="acta_de_desincorporacion"/>
  6 <string key="lifecycle:transition" value="start"/>
</event>
</trace>
  
```

Fig. 2 Ejemplo de un registro de eventos del proceso “Desincorporar productos” Fuente: Elaboración propia.

- **Línea 1** concept:name; muestra el identificador de la instancia del proceso.
- **Línea 2** description; muestra la descripción del proceso.
- **Línea 3** org:resource; muestra quien fue el que ejecuto la actividad.
- **Línea 4** time:timestamp; muestra la fecha y hora en que se ejecutó la actividad.
- **Línea 5** concept:name; muestra lo que se pretende hacer.
- **Línea 6** lifecycle:transition; muestra el ciclo de vida de la transición, en este caso puede tomar dos valores (start, complete).

#### Medidas y métricas

Según una investigación de Chamorro y Maturana [19] como algunas herramientas (*plugin*) que tiene ProM permiten ser configuradas por ejemplo Fuzzy Miner, recomiendan que de no conocerse cabalmente lo que implica cada una de las medidas y métricas, se usen las configuraciones por defecto.

Utilizar por defecto todas las medidas y métricas de la técnica Fuzzy Miner de ProM en su personalización para el HIS garantiza que el modelo generado no sufra importantes variaciones. Si son desactivadas las métricas unarias no se muestra el comportamiento de las actividades del registro de eventos. Si se desactivan las métricas binarias se obtiene un modelo difuso sin representar el comportamiento de las relaciones de precedencia entre actividades. En caso de desactivar las métricas de correlación binaria el modelo se mantiene como el caso de las métricas binarias, debido a que la técnica hace tanto uso de las medidas de significación como las de correlación para generar las relaciones, de no contar con ambas, no se muestran las relaciones entre las actividades.

Considerando lo anterior se puede afirmar que en el caso que se descarte en la personalización cualquiera de las métricas de la técnica Fuzzy Miner, no se mostrarán todas las actividades con sus relaciones limitando los análisis que se hagan sobre los modelos que se generan.

En la personalización de la técnica Fuzzy Miner para el HIS se necesita obtener toda la información posible a partir de los datos que se guardan, por tanto se usan todas las métricas con sus valores por defecto con el objetivo de generar todas las actividades y sus relaciones.

Los filtros originales no se personalizan, el usuario final para hacer uso de ellos necesitará tener conocimientos de lo que significa cada uno, sin embargo, se crean dos filtros el "Frecuencia de actividades" y "Frecuencia de aristas", similar al filtro de Disco [20] para el mismo tipo de análisis. El primero permite mostrar las actividades dependiendo del porcentaje que se quiera visualizar y responde al valor de las métricas unarias. Mientras que el segundo permite mostrar las relaciones de precedencia de dos actividades dependiendo del porcentaje que se quiera visualizar y responde a las métricas binarias y la correlación.

### Fase 2: Integrar la técnica Fuzzy Miner al HIS

En la Figura 3 se muestra el diagrama de clases donde se encuentran las clases necesarias para integrar el componente desarrollado al HIS.

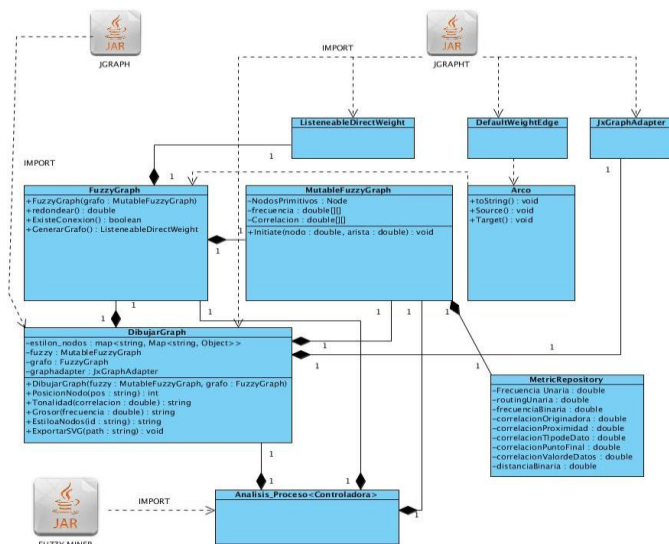


Fig. 3 Diagrama de clases. Fuente: Elaboración propia.

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

### Fase 3: Mostrar el modelo desde el HIS

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

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

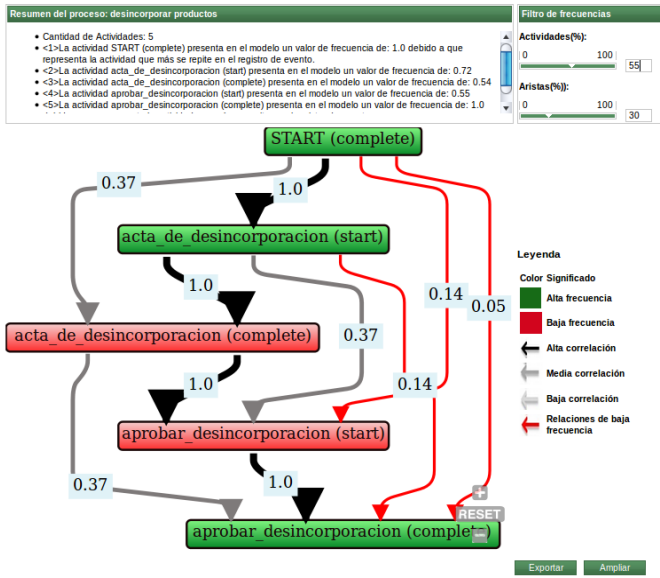


Fig. 4 Ejemplo de proceso modelado utilizando la personalización de Fuzzy Miner en el HIS. Fuente: Elaboración propia.

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

En la parte central se muestra el modelo con la leyenda que provee una explicación de los elementos que intervienen en el mismo, los valores sobre los arcos representan la frecuencia de la relación de las actividades y en la parte inferior se cuenta con la opción de exportar el modelo a formato \*.png para su posible análisis en otros entornos y conjuntamente posee la opción para ampliar el modelo que permite aislarlo a una nueva pestaña del navegador, estas opciones forman parte de los requisitos del componente que genera el registro de eventos.

### III. ANÁLISIS DE PROCESOS DEL XAVIA HIS

#### *Análisis del proceso Desincorporar producto del módulo Almacén*

La Figura 5 muestra un modelo Fuzzy del proceso “Desincorporar producto” del módulo Almacén del HIS. El modelo que contiene los caminos o actividades más repetidas (secuencias principales) en el flujo de información, incluye además las siguientes actividades en su flujo esencial: “START (complete)”, “acta\_de\_desincorporacion (start)”, “acta\_de\_desincorporacion (complete)”, “aprobar\_desincorporacion (start)” y “aprobar\_desincorporacion (complete)”.

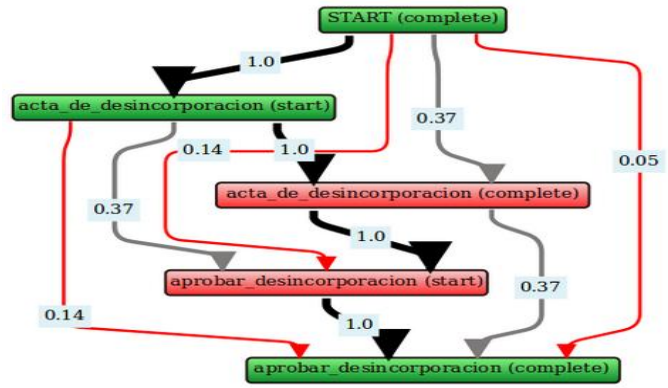


Fig. 4 Secuencia principal del proceso “Desincorporar producto” modelado con con Fuzzy Miner. Fuente: Elaboración propia.

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

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

#### *Análisis del proceso Solicitar producto del módulo Almacén*

Como objeto de análisis se toma el proceso “Solicitar productos” del módulo almacén que cuenta con 15 actividades diferentes y un total de 5713 ocurrencias. El registro de eventos recopila 1140 ejecuciones del proceso en las fechas comprendidas entre 1 de mayo del 2011 y 4 de abril del 2015. La Figura 5 muestra el modelo obtenido.

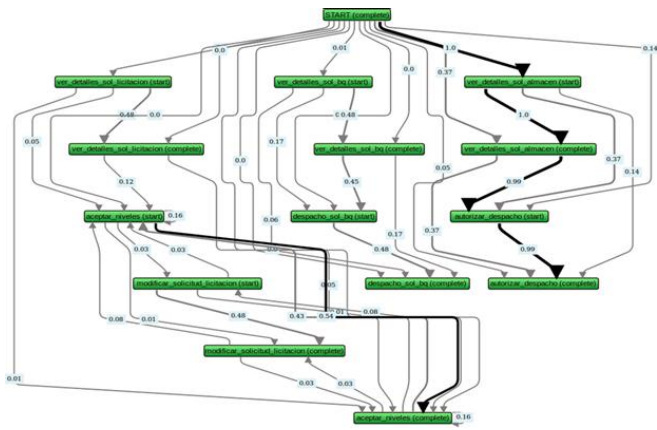


Fig. 5 Modelo del proceso “Solicitar producto” con Fuzzy Miner del HIS. Fuente: Elaboración propia.

Se modela en la herramienta ProM el mismo registro de eventos obtenido. Las actividades de baja frecuencia y alta correlación son agrupadas en un *cluster* por la técnica Fuzzy Miner, así como las de baja frecuencia y baja correlación se eliminan del modelo, el umbral de frecuencia usado aleatoriamente es 55 por ciento. Estos elementos son importantes destacarlos en el modelo para analizarlos, para ello las actividades de baja frecuencia según los criterios definidos por los usuarios se representan en color rojo con la personalización de Fuzzy Miner desde el HIS. El modelo obtenido con ProM se presenta en la Figura 6.

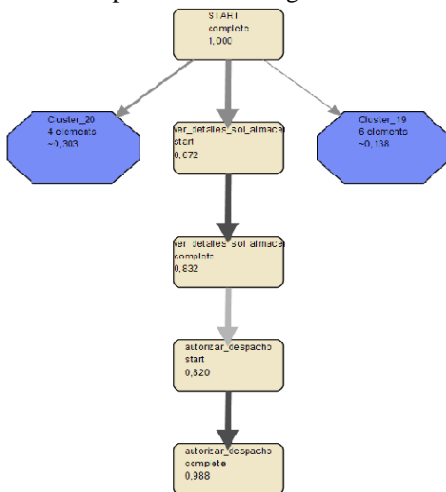


Fig. 5 Modelo del proceso “Solicitar producto” con Fuzzy Miner de ProM. Fuente: Elaboración propia.

En la Figura 6 se muestra el modelo generado por la personalización de la técnica Fuzzy Miner desde el HIS para el valor de frecuencia 55 por ciento con que fue generado el modelo desde el ProM y se destacan en color rojo las 10 actividades que se agrupan en los *clusters*, en este caso no se descartan actividades por lo que se puede afirmar que la personalización propuesta identificó correctamente los elementos de baja frecuencia. Una característica notable del

modelo es la presencia de lazos en las actividades despacho\_sol\_bq (despacho solicitud bloque quirúrgico), modificar\_pedido\_bq (modificar pedido bloque quirúrgico) y aceptar\_niveles. La actividad menos realizada es modificar\_solicitud\_licitacion y la más frecuente es ver\_detalle\_solicitud\_almacen.

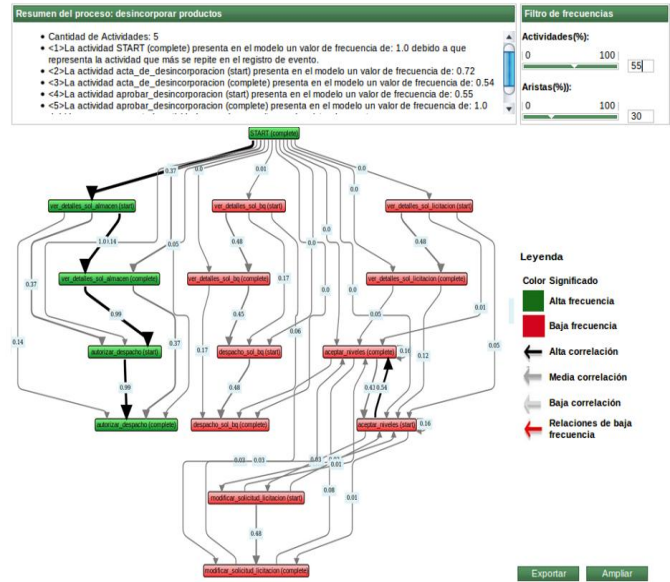


Fig. 6 Modelo del proceso “Solicitar producto” con filtro de frecuencia al 55 por ciento. Fuente: Elaboración propia.

La Tabla 1 presenta un resumen de los análisis realizados sobre modelos obtenidos de siete procesos para generarlo se estableció un valor de frecuencia como indica la columna Valor de frecuencia.

TABLA I  
RESUMEN DEL ANÁLISIS DE FRECUENCIA DE 7 PROCESOS DEL HIS

Procesos del HIS	Rango de fechas	Valor de frecuencia VF	HIS actividades con valor inferior al (VF)
Distribuir producto	2011-05-01 / 2015-05-01	80%	Autorizar_despacho
Desincorporar productos	2011-03-01 / 2015-05-01	42%	Acta_de_desincorporacion Aprobar_desincorporacion
Procesar solicitudes	2012-01-01 / 2015-04-01	85%	crear_contrato(start) crear_contrato(complete) crear_cronograma_de_entrega(start) crear_cronograma_de_entrega(complete)
Transferencia hospitalaria	2011-10-05 / 2015-11-01	75%	Procesar_transferencia_hospitalaria(start)
Interconsultas	2008-01-03 / 2015-10-01	45%	)
Solicitud interconsulta hospitalaria	2011-02-01 / 2015-05-05	70%	Crear_hoja_interconsulta_hospitalaria(start) Crear_hoja_preeanestesia(start)

*Análisis de frecuencia de aristas*

Las tablas 2 y 3 contienen los valores para medir la frecuencia y correlación de aristas en la personalización de la técnica.

TABLA II  
VALORES PARA MEDIR FRECUENCIA DE ARISTAS

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

TABLA II  
VALORES PARA MEDIR LA CORRELACIÓN DE ARISTAS

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

Estos valores facilitan la interpretación de los modelos que se analicen. Propician a los administrativos y analistas de procesos de las instituciones sanitarias instrumentos para una mejor toma de decisiones. A continuación se analizan los procesos “Solicitar producto” y “Desincorporar producto”.

**Proceso Solicitar producto**

Se estableció el filtro de frecuencia de aristas a un 30 por ciento (ver figura 7) aleatoriamente para identificar que actividades tienen baja relación de dependencia. El resultado obtenido al modelar el proceso Solicitar producto se muestra en el modelo de la figura 8.

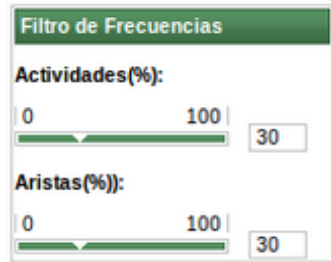


Fig. 7 Interfaz del HIS para la configuración del valor Frecuencia. Fuente: Elaboración propia.

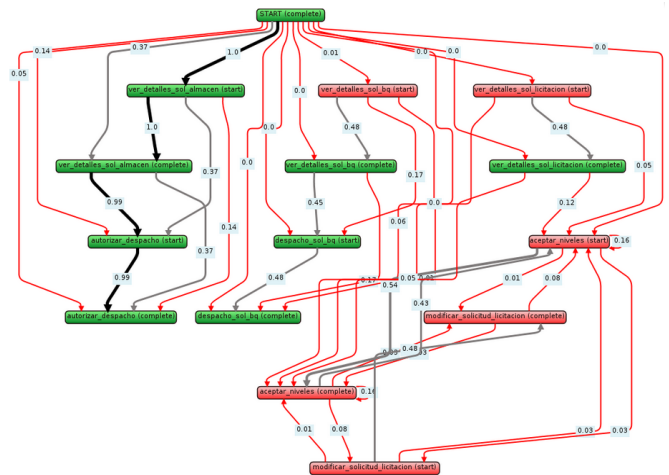


Fig. 8 Modelo con filtro de frecuencia de ejecución menor al 30 por ciento. Fuente: Elaboración propia.

**Proceso Desincorporar producto**

El proceso “Desincorporar producto” del módulo almacén del HIS cuenta con cinco actividades diferentes con un total de 16 ocurrencias. El registro de eventos recopila seis ejecuciones del proceso en las fechas comprendidas entre el 1 de mayo del 2011 y 12 de abril del 2015. El modelo obtenido es representado en la figura 9. Se estableció como valor del Filtro de aristas el 30 por ciento, arrojando como resultado que las relaciones entre las actividades detectadas con valor inferior al del filtro de aristas son destacadas en el modelo resultante en color rojo.

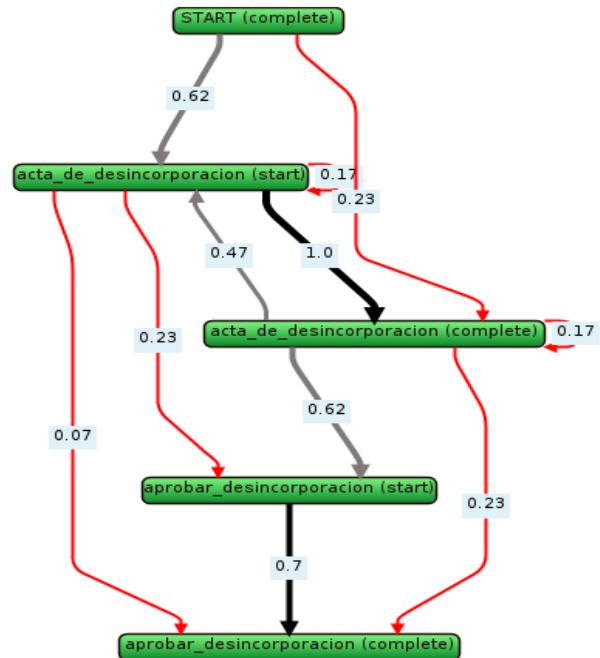


Fig. 9 Interfaz del HIS de aristas con ruido en el proceso “Desincorporar producto”. Fuente: Elaboración propia.

CONCLUSIONES



El estudio de la minería de procesos demuestra que constituye una herramienta de mejora para los centros de salud e instituciones hospitalarias. Estos centros tienden a adoptar una visión de sus procesos por lo que es indispensable un correcto modelado e implementación de los mismos.

Los análisis de procesos en la práctica diaria están centrados en responder cómo debería hacerse y no en cómo se hace realmente, por lo que dichos modelos en ocasiones se generan pareciéndose poco o nada a como realmente se está desarrollando la actividad.

La técnica Fuzzy Miner se ha puesto en práctica en el sector de la salud debido a las gráficas, modelos y secuencias que genera. A partir de su aplicación es posible observar los pasos seguidos por los pacientes, conocer las excepciones, analizar la frecuencia de ejecución e identificar desviaciones en el flujo de los procesos.

La personalización de la técnica Fuzzy Miner se desarrolló basándose en la orientación hacia el Sistema de Información Hospitalaria XAVIA HIS lo cual posibilita la construcción de modelos de acuerdo a las necesidades del personal médico y administrativo, de forma rápida e intuitiva.

Las pruebas realizadas demuestran la necesidad de tener en cuenta todas las métricas de la técnica para su personalización en el Sistema de Información Hospitalaria si se quieren generar modelos completos y con todos los aspectos de interés.

La personalización de la técnica Fuzzy Miner en el Sistema de Información Hospitalaria del CESIM permitió detectar actividades de baja frecuencia en el registro de eventos. Además posibilitó utilizar la técnica desde el sistema, durante la etapa de diagnóstico de procesos.

La aplicación de la personalización del Fuzzy Miner sobre un registro de eventos del HIS permitió realizar análisis de los modelos obtenidos, apoyar investigaciones asociadas a la ejecución de procesos hospitalarios. También fue posible utilizar la información almacenada en el sistema para generar conocimientos.

#### REFERENCES

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

[2] Orellana, A. Herramienta para la detección de variabilidad en procesos hospitalarios. Tesis de Maestría. Universidad de las Ciencias Informáticas, La Habana. 2015.

[3] Orellana, A., Pérez, D., & Larrea., O. U. Analysis of Hospital Processes with Process Mining Techniques. [ed.] Indra Neil Sharkar, Andrew Georgio y Paulo Mazzoncini de Azevedo. Sao Paulo: s.n., 2015. MEDINFO 2015: EHealth-enabled Health: Proceedings of the 15th World Congress on Health and Biomedical Informatics. 2015, Vol. 216, p. 310-314. 978-161449-564-7

[4] Orellana, A., Sánchez Y. & González, L. Aplicación del Modelo L\* de minería de proceso al módulo Almacén del Sistema de

Información Hospitalaria alas HIS. Engineering Education Facing the Grand Challenges, What Are We Doing? Proceedings of the 13th Latin American and Caribbean Conference for Engineering and Technology. PRF#10. 2015. 13 978-0-9822896-8-6.

[5] Orellana, A., Ledesma Y. ¿Por qué aplicar Minería de Proceso en el entorno hospitalario? Engineering Education Facing the Grand Challenges, What Are We Doing? Proceedings of the 13th Latin American and Caribbean Conference for Engineering and Technology, PRF#10, 2015. 13 978-0-9822896-8-6.

[6] Van Doreman, B. Process Mining in Healthcare Systems: An Evaluation and Refinement of a Methodology. Netherlands: Eindhoven University of Technology. 2012.

[7] Salas, A.V. and Rozinat, A. Minería de Procesos. Revista de la Asociación de Técnicos de Informática, 2013, vol. 223, p. 1-55.

[8] Van der Aalst, W.M.P.. Process Mining. Discovery, Conformance and Enhancement of Business Processes. Springer Heidelberg Dordrecht. London New York: Springer. 2011. ISBN 978-3-642-19345-3.

[9] Mans, R.S., Schonenberg, M.H., Song, M., van der Aalst, W.M.P. & Bakker, P.J.M. Process mining in healthcare. A Case Study. Eindhoven University of Technology, Netherlands. 2008.

[10] Mans, R.S., Schonenberg, M.H., Song, M., van der Aalst, W.M.P. & Bakker, P.J.M. Application of Process Mining in Healthcare. A Case Study in a Dutch Hospital, Eindhoven University of Technology. The Netherlands. 2009, p. 425-438.

[11] Günther, C.W. Process Mining in Flexible Environments. Tesis Doctoral. The Netherlands: Eindhoven University of Technology. 2009.

[12] Orellana A., Alfonso, D., and Larrea, O.U. "Generador de Registros de Eventos para el análisis de procesos en el Sistema de Información Hospitalaria xavia HIS." Convención Salud 2015. 2015. ISBN: 978-959-212-963-4,.

[13] van der Aalst, W.M.P. "ProM: The Process Mining Toolkit," in Proceedings of BPM. Germany. Disponible en: CEUR-WS.org. 2009. Vol. 489, s.n.

[14] Orellana A., Pérez, Y. E. and Larrea, O.U. "Process Mining in Healthcare: Analysis and Modeling of Processes in the Emergency Area." Latin America Transactions, IEEE (Revista IEEE America Latina). 2015. ISSN: 1548-0992. 13:5.

[15] van der Aalst, W.M.P. y Günther, C.W. Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics. Business Process Management. Alonso, G., Dadam, P. y Rosemann, M. Berlín, Alemania: Springer- Berlin Heidelberg, 2007. p. 328–343. ISBN 978-3-540-75182-3.

[16] Process Mining Group. Fuzzy Minner-como usar el plugin en ProM. [en línea]. [Consulta: 24 febrero 2015]. Disponible en: <http://www.processmining.org/online/fuzzyminer>. 2009.

[17] LI, Jin, et al. Fuzzy keyword search over encrypted data in cloud computing. En INFOCOM, 2010 Proceedings IEEE. IEEE, 2010. p. 1-5.

[18] De Weerd, J. et al. A multi-dimensional quality assessment of state-of-the-art process discovery algorithms using real life event logs. Information Systems., 2012, vol. 37, p. 654-676.

[19] Chamorro, M.C. y Maturana, S. Método para Aplicar Minería de Procesos a la Distribución de Bebestibles No Alcohólicos. Tesis de Maestría. Chile. 2013.

[20] Fluxicon. Process mining for professionals. Discover your process. [Online] Fluxicon, 2015. [Cited: Noviembre 30, 2015.] <https://fluxicon.com/disco/>.2015.