

Review of Models for Applying Process Mining Techniques

Arturo Orellana García

aorellana@uci.cu

Universidad de las Ciencias Informáticas

Damián Pérez Alfonso

dalfonso@uci.cu

Universidad de las Ciencias Informáticas

Vivian Estrada Sentí

vivian@uci.cu

Universidad de las Ciencias Informáticas

ABSTRACT

Process mining is a novel alternative to improve processes in a variety of application domains. It aims to extract information from data stored in records of traces from information systems, looking for errors, inconsistencies, vulnerabilities and variability in processes that are executing. The process mining techniques are used in multiple sectors such as industry, web services, business intelligence and health. However, to apply these techniques there are several models and little information on which of them is better to apply, by not having a comparative analysis of these. The research focuses on collecting information on models proposed by author's worldwide reference in the process mining topic, to apply techniques for the discovery, conformance checking and process improvement. Is performed a brief analysis of them in order to select the most comprehensive for its application in health environment. Current research contributes to the development of a model for the detection and analysis of variability in hospital processes using process mining techniques. Allows to readers to have in a centrally way, criteria for deciding which model to use, or what steps to employ of one or more models.

KEYWORDS:

Analysis, conformance checking, discovery, improvement, model, process mining, variability.

Revisión de los principales modelos para aplicar técnicas de Minería de Procesos

Arturo Orellana García

aorellana@uci.cu

Universidad de las Ciencias Informáticas

Damián Pérez Alfonso

dalfonso@uci.cu

Universidad de las Ciencias Informáticas

Vivian Estrada Sentí

vivian@uci.cu

Universidad de las Ciencias Informáticas

RESUMEN

La minería de procesos constituye una alternativa novedosa para mejorar los procesos en una variedad de dominios de aplicación. Tiene como objetivo extraer información a partir de los datos almacenados en los registros de trazas de los sistemas de información, en busca de errores inconsistencias, vulnerabilidades y variabilidad en los procesos que se ejecutan. Las técnicas de minería de procesos se utilizan en múltiples sectores, como la industria, los servicios web, la inteligencia de negocios y la salud. Sin embargo, para aplicar estas técnicas existen varios modelos a seguir y poca información sobre cual aplicar, al no contar con un análisis comparativo entre estos. La investigación se centró en recopilar información sobre los principales modelos propuestos por autores de referencia mundial en el tema de minería de procesos para aplicar técnicas en el descubrimiento, chequeo de conformidad y mejoramiento de los procesos. Se realiza un análisis de los mismos en función de seleccionar los elementos y características útiles para su aplicación en el entorno hospitalario. La actual investigación contribuye al desarrollo de un modelo para la detección y análisis de variabilidad en procesos hospitalarios utilizando técnicas de minería de procesos. Permite a los investigadores tener de forma centralizada, los criterios para decidir qué modelo utilizar, o qué fases emplear de uno o más modelos.

PALABRAS CLAVE: Análisis, chequeo de conformidad, descubrimiento, mejoramiento, minería de proceso, modelo, variabilidad.

INTRODUCCIÓN

El significativo avance de las Tecnologías de la Información y la Comunicación (TIC) a nivel mundial, ha propiciado el desarrollo informático de las organizaciones. En la actualidad, la mayoría de las tareas de oficina, en las empresas o incluso en lugares como bancos, hospitales, centros educativos y agencias de viajes, son ejecutadas y controladas

por sistemas de información con el objetivo de recolectar, procesar, almacenar y distribuir datos en apoyo al control y la toma de decisiones.

Las organizaciones actuales reconocen la ventaja que supone incorporar los nuevos avances tecnológicos a sus procesos de negocio. 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.

Deming plantea que la mejora de la calidad de los procesos implica mayor productividad (Deming, 1989), por tanto se hace necesario romper viejos paradigmas e incorporar nuevos para optimizar los procesos en las organizaciones. Se convierte en el aspecto de mayor importancia el análisis de los datos que se generan durante la ejecución de los procesos.

Mediante el enfoque basado en procesos se pueden identificar y gestionar un gran número de procesos relacionados entre sí y analizar y seguir de forma racional su desarrollo. Permite además, obtener la mejora continua de los resultados por medio de la eliminación de errores y procesos redundantes en las diferentes funciones de la organización. El aumento de la complejidad de este enfoque, enmarcado por el crecimiento del volumen de datos que se almacena continuamente, ha favorecido el surgimiento y desarrollo de tecnologías inteligentes, convirtiendo estos datos en información relevante para la toma de decisiones.

Las técnicas de minería de proceso, permiten extraer información no trivial y útil de los registros de evento 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 negocio. Su objetivo fundamental es descubrir, monitorear y mejorar los procesos reales (van der Aalst et al., 2011) a través de la extracción de conocimiento de los registros de eventos ampliamente disponibles en los actuales sistemas de información.

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, 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.

La minería de proceso se propone como una alternativa novedosa para mejorar los procesos en una variedad de dominios de aplicación, desde software de aplicaciones, sistemas de inteligencia artificial hasta aplicaciones web. Es una disciplina de investigación relativamente joven que se ubica entre la inteligencia computacional y la minería de datos, por una parte; y la modelación y análisis de procesos, por otra (van der Aalst, 2011). Surge con el objetivo de analizar los registros de eventos destinados a recopilar las trazas de los procesos que ocurren en los sistemas de información actuales, en busca de errores, inconsistencias, vulnerabilidades, desvíos y variabilidad.

La variabilidad en la ejecución de los procesos de negocio (La Rosa, et al., 2013) influye directamente en la productividad y la calidad de los servicios. Una esfera social afectada directamente es la Salud. Se ha calculado que la industria de la salud ha incurrido en el deceso de más de 400.000 pacientes al año sólo en EEUU, a partir de los errores médicos

evitables (James, 2013). Muchas de estas muertes evitables se deben a procesos ineficaces e ineficientes, y los traspasos entre actividades y procesos.

Mathews Loxton (Loxton, 2015), investigador de procesos hospitalarios plantea que " (...) las investigaciones orientadas a las acciones sobre la variabilidad en procesos de salud constituyen una prioridad, donde los métodos y herramientas de análisis de datos pueden desempeñar un papel importante y fundamental." La identificación de los procesos y la reducción de la variabilidad se traducen en una mejora significativa en los índices de morbilidad y mortalidad, el ahorro de recursos y control de las actividades. Esta situación condujo a la necesidad de realizar una revisión de los principales modelos para aplicar técnicas de minería de proceso.

El presente artículo tiene como objetivo realizar un análisis documental de los principales modelos utilizados para guiar proyectos de minería de procesos; en función de obtener los criterios necesarios para el desarrollo de un modelo que adapte e integre técnicas de minería de procesos para detectar variabilidad en la ejecución de procesos hospitalarios.

METODOLOGÍA

Definición de modelo de proceso

Antes de definir el modelado de procesos de negocio, el término "modelado" tiene que ser discutido en un contexto más general. Nordsieck ha hecho hincapié en que "la utilización de los símbolos permite al modelo no sólo reemplazar o complementar el lenguaje natural para la representación de asuntos complejos, sino también, mostrar la noción del objeto, a menudo de una manera más integral, como con cualquier otra forma de representación" (Nordsieck, 1932). Las características más protuberantes de un modelo son la brevedad, claridad, precisión, y su calidad gráfica (Nordsieck, 1932). Stachowiak define un modelo como el resultado de una cartografía de la simplificación de la realidad que sirve a un propósito específico (Stachowiak, 1973). De acuerdo con esta percepción, hay tres cualidades importantes que un modelo debe poseer.

En primer lugar, hay una asignación que establece una representación de los originales naturales o artificiales que pueden ser modelos por sí mismos. En segundo lugar, sólo los atributos del original que se consideran relevantes se asignan al modelo; el resto se omite. Por lo tanto, el modelo proporciona una abstracción en términos de un homomorfismo en un sentido matemático (Kühne, 2006). En tercer lugar, el modelo es utilizado por el modelador en lugar del original en un cierto punto en el tiempo y para un fin determinado. Esto significa que un modelo siempre implica la pragmática.

Las Directrices de Modelado (GoM, por sus siglas en inglés) (Becker y otros, 1995), (Schütte y Rotthowe, 1998), (Becker y otros, 2000) definen los principios que sirven a este propósito la normalización. Son aplicables para cualquiera de las posiciones epistemológicas, el positivismo y el constructivismo, debido a que tanto la elección de un determinado homomorfismo (posición positivista) y la percepción del modelador (posición constructivista) introducen elementos subjetivos. Por lo tanto, las Directrices de Modelado (GoM) (Becker y otros, 1995), (Schütte y Rotthowe, 1998) incluyen seis principios concretos para el logro de la intersubjetividad de modelos. Los tres primeros definen las condiciones previas necesarias para la calidad de los modelos, es decir, la exactitud,

pertinencia y eficiencia económica, y los otros tres son opcionales, es decir, claridad, comparabilidad y diseño sistemático.

Corrección: En primer lugar, un modelo tiene que ser sintácticamente correcto. Este requisito exige el uso de primitivas de modelado permitidas y la combinación de ellas de acuerdo a reglas predefinidas. En segundo lugar, el modelo debe ser semánticamente correcto. Por lo tanto, tiene que ser formalmente correcto y consistente con el (percepción del) mundo real.

Pertinencia: Este criterio exige que sólo partes interesantes del universo de discurso se reflejan en el modelo. Es, por lo tanto, en relación con la noción de integridad como se propone en (Batini y otros, 1986).

Eficiencia económica: Esta guía presenta un equilibrio entre los beneficios y los costos de poner los otros criterios a la práctica. Por ejemplo, la corrección semántica podría ser despreciada en cierta medida, en caso de lograrlo es costosa.

Claridad: Esta es una guía muy subjetiva exigiendo que el modelo debe ser entendido por parte del usuario del modelo. Está principalmente relacionada con las convenciones de diseño o la complejidad del modelo.

Comparabilidad: Este criterio exige la utilización coherente de un conjunto de directrices en un proyecto de modelado. Entre otros, se refiere a los convenios de nomenclatura.

Sistemática de diseño: Esta directriz exige una separación clara entre los modelos en diferentes puntos de vista (por ejemplo, los aspectos estáticos y aspectos conductuales) y mecanismos definidos para integrarlos.

Principales modelos para aplicar técnicas de minería de proceso

En esta sección serán descritos algunos de los principales modelos propuestos por autores referenciados dentro del campo de la minería de proceso, con el objetivo de analizar su funcionamiento básico. De forma general, se presentan las etapas, fases o pasos dentro de los tipos de minería de proceso descubrimiento, conformidad y extensión o mejora.

Modelo de van Giessel

Este modelo (van Giessel, 2004) propone aplicar minería de proceso en un sistema SAP (Sistemas, Aplicaciones y Productos en Procesamiento de datos). Está compuesto por dos fases fundamentales y cinco pasos.

Fase 1: se determina cuáles tablas son relevantes a través de los objetos de negocio. Se establece un enlace entre los objetos de negocio y el proceso de negocio, mediante componentes de la aplicación. Esta fase está parcialmente automatizada por la herramienta TableFinder.

Fase 2: se recupera el flujo del documento. Primeramente se localizan los datos de las tablas y se extraen. La recuperación se realiza mediante la agrupación de todos los datos que tienen el mismo número de documento o que están relacionadas unas con otras. Cuando todos los datos relacionados se agrupan, se recupera el flujo de documentos que al final se exporta en un archivo XML.

Una vez terminadas estas dos fases se procede a aplicar técnicas de minería a través de los siguientes pasos:

Paso 1: determinar el modelo del proceso. Se descubre el modelo del proceso que representa las actividades ejecutadas por los usuarios. Este modelo del proceso puede ser comparado con el proceso de negocios definido, utilizando el análisis-delta. Todas las actividades ejecutadas se monitorean y la información se utiliza para determinar el modelo del proceso. Existen herramientas como EMiT y Little Thumb para generar estos modelos, la primera centrada en el tiempo y la segunda en el ruido.

Paso 2: análisis de la red social. Se centra en las relaciones entre los individuos o grupos de personas; mediante la herramienta MISON se puede descubrir las relaciones y realizar el análisis. Las relaciones se construyen por métricas predefinidas y los resultados se pueden mostrar en la representación de una matriz o en un gráfico.

Paso 3: determinar los parámetros básicos de rendimiento. La herramienta ARIS Process Performance Manager, calcula los indicadores pre-definidos de rendimiento de procesos, pero también se pueden establecer indicadores de rendimiento de procesos a sí mismo. La minería de proceso calcula los indicadores básicos de rendimiento del proceso a partir del registro de eventos.

Paso 4: determinar el modelo de la organización. Las empresas están organizadas de diversas formas tanto funcional, lineal como multi-funcional, donde cada integrante es responsable de actividades específicas. Si se desea obtener el modelo de organización sin previo conocimiento acerca de la estructura real, se puede utilizar minería de proceso, ya que en el registro de eventos se guarda el originador de cada instancia de un proceso, a través del cual se puede determinar qué actividades son ejecutadas por quiénes y por tanto, obtener el modelo de la organización.

Paso 5: análisis de las características de rendimiento. Este paso se basa en el "si-entonces", ya que cada instancia de proceso tiene características específicas y a través de las técnicas de minería se pueden encontrar relaciones significativas entre diferentes características de rendimiento.

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. Permite visualizar las dependencias entre las técnicas de minería de procesos a utilizar, su representación gráfica puede ser consultada en (Rozinat et al., 2009). A continuación se describen los pasos que componen el modelo.

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 desean 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.

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 el Fuzzy Miner, (plugin de la herramienta ProM), 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.

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).

Está compuesto por cinco etapas que se describen a continuación:

Etapas 0: planificar y justificar. Los proyectos de minería de proceso 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 proceso:

- 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 proceso. 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.

Caso de estudio de Jans

Este caso de estudio no se encuentra estandarizado como modelo, sin embargo constituye un ejemplo para guiar un proyecto de minería de procesos (Jans, 2011). A continuación se

realiza un resumen de los pasos fundamentales como resultado de la aplicación del caso de estudio:

Paso 1: caracterización del proceso. Para poder obtener buenos resultados en el proceso de minería es necesario realizar un exhaustivo análisis del proceso, en el que se recogen todos los datos relevantes y se crea un mapa de dicho proceso. Este mapa consta de cuatro componentes: objetivos del proceso, actividades, flujos de información y contabilidad del impacto, siendo el segundo componente el más importante por constituir la entrada principal al Paso 2, o sea, la creación del registro de eventos.

Además se realizan consultas a los ejecutivos de negocios y especialistas del sistema de información, entrevistas y observaciones a los empleados de distintos departamentos durante horario laboral; así como la revisión de los manuales de usuarios internos del sistema.

Paso 2: creación del registro de eventos. Luego de realizar toda la caracterización del proceso se procede a construir el registro de eventos adecuado, a partir de los datos capturados por el sistema de información y almacenados en el registro. Estos datos se encuentran en grandes cantidades y dispersos por numerosas tablas del sistema, las cuales son específicas para cada sistema y cada proceso de negocio. Para minar los datos es necesario configurarlos dentro del registro de eventos con unos requisitos de formato que facilite la minería de proceso.

Este formato constituye la estructura de los datos relevantes alrededor de las actividades que conforman el proceso. Para crear el registro de eventos es necesario conocer cuáles son las actividades que constituyen el proceso y cuál es el caso o instancia de proceso que es el foco de estas actividades, a través de lo que se define el contenido y la estructura de este. Luego es necesario convertir el registro de eventos en el formato MXML.

Paso 2.1: refinamiento del registro de eventos. Una actividad, un caso, un usuario y el tiempo son las características para minar un proceso. El momento en que las actividades y el originador son ejecutados, son creadas meta data por el sistema de información y no se basan en información de entrada creada por el usuario. Esta información de entrada que se contraponen a la meta data es llamada atributos.

Paso 3: análisis con minería de proceso. Después de crear el registro de eventos, se analiza el proceso a través de las siguientes tareas:

Tarea 1: descubrimiento de proceso. Se analiza el orden de las actividades para descubrir cómo se efectúa el proceso en la vida real.

Tarea 2: análisis de roles. Se analiza la asociación de las actividades con las personas a través de la prueba de controles de segregación de tareas.

Tarea 3: verificación por análisis de atributos. Se analiza además la salida de los pasos previos. Se utilizan los atributos de los casos con el fin de verificar la eficacia de los controles internos de la empresa a lo largo del proceso en análisis.

Tarea 4: análisis de redes sociales. Se realiza una búsqueda de las redes sociales entre los empleados que participan en el proceso, con el objetivo de comprender mejor la interacción entre las funciones y actividades para las transacciones anómalas.

En resumen, se puede determinar el grado de concordancia entre lo detectado por los auditores internos en su análisis estándar y la información descubierta al aplicar el proceso de minado.

RESULTADOS

Los modelos descritos en la sección precedente tienen en común la forma de guiar el proyecto de minería. La Figura 1 representa el flujo de este proceso y los elementos necesarios para obtener resultados satisfactorios.



Figura 1: Modelo conceptual del proceso de minería. Fuente: Elaboración propia.

Para aplicar minería de procesos es necesario el uso de un conjunto de herramientas orientadas a esta finalidad. El proceso de minería comienza utilizando la herramienta XESame (Verbeek et al., 2011), Eventifier (Rodríguez et al., 2012) o un equivalente para extraer la información necesaria de la bitácora del sistema en análisis. Para esto es necesario tener conocimientos sobre la base de datos del sistema, las configuraciones de las herramientas de extracción y las características que deben cumplir las salidas. Como resultado de esta parte del proceso se genera un registro de eventos y se exporta en formato estándar XES (eXtensible Event Stream).

El registro de eventos es importado a herramientas de análisis como ProM (van der Aalst et al., 2009) o Disco (Fluxicon, 2009), la primera es una herramienta académica de libre acceso y la segunda es comercial, lo cual dificulta su adquisición. El marco de trabajo ProM provee más de 600 técnicas con capacidades de minería de procesos, una cifra que dificulta saber cuáles usar en determinados contextos o situaciones, además de que cada técnica tiene un conjunto de métricas y parámetros de necesaria configuración. Como resultado se obtiene un modelo de la ejecución del proceso en análisis.

Las herramientas de minería de procesos, aunque son eficientes en su campo, poseen alta complejidad configuraciones técnicas y procedimientos engorrosos, tanto para extraer registros de eventos como para obtener modelos. Por lo que la aplicación de esta tecnología en su estado actual es limitada para profesionales no expertos.

Luego de analizar cada uno de los modelos anteriormente descritos, se resumió en la Tabla 1 las principales características que los definen.

Tabla 1: Comparación de modelos sobre minería de proceso.

Características	Modelos				
	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	✓				✓
Utilización de técnicas	✓	✓	✓	✓	✓
Análisis de rendimiento		✓	✓	✓	✓
Análisis de roles	✓	✓	✓	✓	
Análisis con el cliente				✓	✓
Análisis desde la perspectiva organizacional	✓	✓	✓	✓	✓
Análisis desde la perspectiva de control de flujo	✓	✓	✓	✓	✓
Análisis desde la perspectiva de tiempo	✓	✓	✓	✓	✓
Análisis desde la perspectiva de caso				✓	✓
Otras perspectivas		✓	✓		✓
Procedimiento para analizar el proceso					✓
Elaborado para un dominio de aplicación		✓			
Retroalimentación de los resultados					✓
Herramientas para el análisis	✓				
Considera profesionales no expertos					
Especificación de las técnicas a utilizar			✓	✓	

Fuente: Elaboración propia.

DISCUSIÓN Y CONCLUSIONES

Tras un análisis de las características resumidas en la Tabla 1 se concluye que los modelos de Jans y Giessel no permiten obtener una caracterización del proceso, a diferencia de Rozinat, Bozkaya y van der Aalst. En los modelos de Giessel, Rozinat y Boskaya no se obtiene un modelo de proceso integrado. El modelo más completo a partir de los parámetros de comparación definidos es el de van der Aalst, sin embargo posee como limitante que la forma de su funcionamiento es orientada a expertos de la minería de procesos. Solo el modelo de van der Aalst presenta un procedimiento para guiar el proyecto de minería de procesos. 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. Brinda además, soporte a cualquiera de los tres tipos de minería de procesos, descubrimiento, conformidad y mejora.

Los modelos en general no se enfocan en el uso de técnicas específicas, sino que permite aplicar una gran variedad de estas, lo cual dificulta la realización de análisis para los no expertos si se considera que existen más de 600 técnicas desarrolladas. Todos los modelos permiten realizar análisis desde una perspectiva organizacional, lo que supone una ventaja para los administrativos, sin embargo no especifican cómo hacerlo. Son modelos elaborados para proyectos de minería genéricos, solo la propuesta de van Giessel se centra en un dominio de aplicación. Un aspecto importante a destacar es que solo el modelo de van der Aalst permite retroalimentar los resultados, lo cual permite cerrar el ciclo del modelo y evaluar el impacto de los resultados.

Se consultó una importante gama bibliográfica, detectándose en la misma, la descentralización de criterios para seleccionar una guía o procedimiento para la obtención de modelos de proceso. La aplicación de uno u otro modelo posibilita tomar decisiones en aras de mejorar los procesos hospitalarios e incidir directamente sobre la gestión de los recursos, la calidad de los servicios a pacientes y el control y distribución del personal sanitario. Se propone combinar características de los modelos descritos para la elaboración de procedimientos desarrollados a la medida en aras de fortalecer los análisis de procesos de negocio.

El modelo de van der Aalst será utilizado para guiar el desarrollo de una herramienta de minería de procesos para detectar variabilidad en procesos hospitalarios. La cual contará con una personalización de algunas técnicas fundamentales para el análisis de los procesos que se gestionan en el sistema antes mencionado.

En función de contribuir a solucionar las deficiencias detectadas se propone como extensión de la presente investigación, la elaboración y aplicación de un modelo que permita detectar y analizar variabilidad en procesos hospitalarios. Para ello es necesario tener en cuenta los siguientes principios:

- ✓ Integración de las técnicas de minería de procesos al modelo, propiciando conocer cuáles utilizar en los análisis.
- ✓ Los usuarios no tienen que conocer las técnicas de minería de procesos, sus configuraciones y métricas, extendiendo su uso a usuarios no expertos en esta tecnología.
- ✓ Implementación de estándares, componentes y procedimientos para detectar de forma sencilla la existencia de variabilidad en la ejecución de procesos sanitarios.
- ✓ Evaluar la calidad de los datos contenidos en los registros de eventos para realizar análisis de frecuencia de ejecución y tiempo, así como visualizar las variantes alternativas de los procesos y la identificación de cuellos de botella.

- ✓ Extensión a sistemas de otros sectores sociales para contribuir a la toma de decisiones sobre la ejecución de sus procesos de negocio.
- ✓ Pertinencia como garantía de la adecuación del modelo en el contexto de la salud.
- ✓ Actualización permanente mediante la retroalimentación de la información que nutre al modelo.

REFERENCIAS

- Batini, C., Lenzerini, M., and Navathe, S. B. (1986). A Comparative Analysis of Methodologies for Database Schema Integration. *ACM Computing Surveys*, pp. 323–364.
- Becker, J., Rosemann, M., and Schütte, R. (1995). Grundsätzeordnungsmässiger Modellierung. *Wirtschaftsinformatik*, p. 435–445.
- Becker, J., Rosemann M., and von Uthmann, C. (2000). Guidelines of Business Process Modeling. In van der Aalst, W.M.P., Desel, j., and Oberweis, A. editors, *Business Process Management. Models, Techniques, and Empirical Studies*, pp. 30–49. Berlin et al.
- Bozkaya, M., Gabriels, J. & van der Werf, J. M. (2009, Febrero). Process Diagnostics: a Method Based on Process Mining. Trabajo presentado en The International Conference on Information, Process, and Knowledge Management, Cancun, México.
- Deming, W. E. (1989). *Calidad, productividad y competitividad: La salida de la crisis*. España: Díaz de Santos S.A.
- Fluxicon. (2015). Fluxicon. Process mining for professionals. Discover your processes. [Online] Fluxicon, 2015. [Cited: Noviembre 30, 2015.] <https://fluxicon.com/disco/>.
- James, J. T. (2013). A new, evidence-based estimate of patient harms associated with hospital care. *Journal of patient safety*, 9(3), 122-128.
- Jans, M. (2011). Process Mining of Event Logs in Internal Auditing: A Case Study. The 2nd International Symposium on Accounting Information Systems, Italy.
- Kühne, T. (2006). Matters of (meta-) modeling. *Software and Systems Modeling*, pp. 369–385.
- La Rosa, M., et al. Business process variability modeling: A survey. 2013.
- Loxton, M. 2015. Using MAXQDA in Healthcare Research: Process Discovery. MAXQDA. The Art of Data Analysis. [En línea] 23 de Noviembre de 2015. [Citado el: 30 de Noviembre de 2015.] <http://www.maxqda.com/process-discovery-maxqda>.
- Nordsieck, F. (1932). Die Schaubildliche Erfassung und Untersuchung der Betriebsorganisation. *Organisation - Eine Schriftenreihe*. C. E. Poeschel Verlag, Stuttgart, p. 23.
- Rodríguez, C., Engel, R., Kostoska, G., Daniel, F., Casati, F., & Aimar, M. (2012, September). Eventifier: Extracting process execution logs from operational databases. In *Demonstration Track of BPM Conference, CEUR-WS* (pp. 17-22).
- Rozinat, A., Mans, R. S., Song, M., & van der Aalst, W.M.P. (2009). Discovering Simulation Models. *Information Systems*, Vol. 34(3), p. 305-327.
- Schütte, R. and Rotthowe, T. (1998). The Guidelines of Modeling - An Approach to Enhance the Quality in Information Models. In Tok Wang Ling, Sudha Ram, and Mong-Li Lee, editors, *Proceedings of the 17th International Conference on Conceptual Modeling*, volume 1507 of *Lecture Notes in Computer Science*, p 240–254, Singapore.
- Stachowiak, H. (1973). *Allgemeine Modelltheorie*. Springer-Verlag.
- 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. (2011). *Process Mining. Discovery, Conformance and Enhancement of Business Processes*. London New York: Springer.

- Van Giessel, M. (2004). Process mining in SAP R/3. Tesis de Maestría, Eindhoven University of Technology, The Netherlands.
- Verbeek, H., Buijs, J., Van Dongen, B., and Van Der Aalst, W. (2011). XES, XESame, and ProM6, volume 72 LNBIP of CAiSE Forum 2010 on Information Systems Evolution. Hammamet.
- 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.

Artículo recibido el 20/11/2014 – publicado el 18/03/2016.