

NEW SOFTWARES/NOVEDADES DE SOFTWARE

ALTERNATIVA PARA REALIZAR ANÁLISIS DE RENDIMIENTO DE PROCESOS AUTOMATIZADOS EMPLEANDO TÉCNICAS DE MINERÍA DE PROCESOS Y SIMULACIÓN SOBRE REDES DE PETRI

Amanda Gramatges Hermida¹, Mayret Suárez Rodríguez, Martha Valdés Padrón, Lisbet Ross Domas, Regina Reyes López.

Complejo de Investigaciones Tecnológicas Integradas, CITI.

Facultad de Ingeniería Industrial, Instituto Superior politécnico José Antonio Echeverría, Cuba

ABSTRACT

Nowadays Process Mining and Simulation are recognized like innovative alternatives for business process analysis and improvement. Traditional techniques of process analysis need generally long processing times and are supported on analysts expertise, with the corresponding risk of human error; however, process mining techniques are able to extract knowledge from event logs commonly available in today's information systems through less expensive methods in time and resources. Whereas traditional simulation approaches rely on hand-made models, we advocate the use of process mining techniques for creating more reliable simulation models based on real event data. The present paper shows the results obtained with process mining and simulation on Petri Nets and ARENA in the same process, in order to compare the results obtained by each of them and determine if there are significant differences between them.

KEYWORDS: Process mining, simulation, CPN Tools, Petri Nets, ProM, ARENA.

MSC: 68N19

RESUMEN

En la actualidad se reconoce la Minería de Procesos y la Simulación como alternativas novedosas soportadas en las tecnologías de la información y las comunicaciones para el análisis y la mejora de procesos. Las técnicas tradicionales para el análisis de los procesos requieren de tiempos relativamente altos y se sustentan sobre la base de la experticia de los analistas, con el correspondiente riesgo a la aparición de error humano; sin embargo, la Minería de Procesos posibilita la recuperación y análisis de información, a partir de registros de eventos, empleando métodos menos costosos en tiempo y recursos; así como, permite realizar análisis adicionales sobre el proceso descubierto como la determinación de patrones de comportamiento y análisis de rendimiento. El propósito del estudio es realizar una comparación entre los métodos tradicionales que utilizan simulación en ARENA para el análisis de los procesos y una nueva alternativa basada en Minería de Procesos y simulación sobre Redes de Petri empleando el simulador CPN Tools, definir concretamente ámbitos de aplicación de cada herramienta y realizar una comparación de los resultados obtenidos con la utilización de cada una de ellas en un caso de estudio.

1. INTRODUCCIÓN

La efectividad, eficacia y eficiencia de los procesos de negocio de una organización son vitales para su supervivencia en el mundo competitivo empresarial del momento. Procesos con diseños inadecuados pueden conducir a tiempos de respuesta prolongados, bajos niveles de servicio, utilización de recursos desbalanceada

¹ E-mail: agramatges@citi.cu

y clientes insatisfechos. Es por esto que es importante analizar los procesos antes de su implementación, para descubrir fallas en el diseño, y durante su ejecución, para obtener un diagnóstico e identificar posibles mejoras [15].

Las técnicas tradicionales para el análisis de los procesos requieren de tiempos relativamente altos y se sustentan sobre la base de la experticia de los analistas, con el correspondiente riesgo a la aparición de error humano; sin embargo, la Minería de Procesos posibilita la recuperación y análisis de información, a partir de los registros de eventos, empleando métodos menos costosos en tiempo y recursos; así como, permite realizar análisis adicionales sobre el proceso descubierto como la determinación de patrones de comportamiento y análisis de rendimiento.

Entre los espacios de conocimiento en los que ha impactado la Minería de Procesos destacan la obtención de la información y la generación automática de modelos de simulación en Redes de Petri. Tradicionalmente la construcción de modelos de simulación se estructuran sobre el análisis y búsqueda de información en una muestra del entorno para realizar inferencias al resto de la población, incurriéndose en riesgos de cometer errores al generalizar los resultados; la Minería de Procesos, unida a la existencia de grandes volúmenes de datos digitales almacenados en los sistemas de información, permite revolucionar el campo de la simulación extendiendo la muestra a toda la población y reduciendo el consumo de recursos humanos y materiales para su ejecución.

Una de las ventajas de la utilización de Redes de Petri es que BPMN, EPCs, UML ADs, y otras muchas notaciones de modelación de procesos de negocio, tienen en común que todas utilizan semántica basada en tokens. Consecuentemente, existen muchas técnicas y herramientas para convertir las Redes de Petri a BPMN, BPEL, EPCs y UML ADs y viceversa. Como resultado, los conceptos fundamentales de las Redes de Petri son utilizados indirectamente para realizar análisis, representar modelos, y clarificar semánticas [15].

El desarrollo de las Redes de Petri, y todas sus extensiones, ha estado motivado por el deseo de crear un lenguaje de modelación capaz de representar procesos industriales y al mismo tiempo teóricamente bien fundado y lo suficientemente versátil para ser utilizado en la práctica por sistemas de gran tamaño y complejidad. Las Redes de Petri y sus extensiones poseen un estándar ISO/IEC [5], [6], [7] y tanto el lenguaje de modelación como las herramientas que lo soportan cumplen con este estándar.

La aplicación práctica de las Redes de Petri radica en la combinación de la simulación automática e interactiva, visualización, y análisis de rendimiento. Estas actividades en conjunto resultan en la validación del sistema objeto de estudio para justificar que posee las propiedades deseadas y se obtiene un alto nivel de confiabilidad y entendimiento de este sistema. Los modelos de Redes de Petri, y sus extensiones, pueden utilizarse para validar la precisión lógica y funcional del sistema, además de su rendimiento [8]. Es por este motivo que existen numerosas aplicaciones de la modelación y simulación en Redes de Petri de procesos de negocio y flujos de trabajo, por solo citar algunas: [8], [11], [12], [13], [14], [15], [4].

Con vistas a potenciar la aplicación de técnicas y herramientas de Minería de Procesos y simulación sobre Redes de Petri en función de la gestión empresarial, se estudian nuevas alternativas que permitan realizar análisis de rendimiento a procesos automatizados empleando técnicas de simulación sobre registros de eventos; y posibiliten el análisis de los procesos de la organización a partir de las trazas generadas por los sistemas informáticos que les dan soporte.

El propósito del estudio es realizar una comparación entre los métodos tradicionales que utilizan simulación en ARENA para el análisis de los procesos y una nueva alternativa basada en Minería de Procesos y simulación sobre Redes de Petri empleando el simulador CPN Tools, definir concretamente ámbitos de aplicación de cada herramienta y realizar una comparación de los resultados obtenidos con la utilización de cada una de ellas en un caso de estudio.

Como apoyo al proceso de investigación se hace uso de las técnicas y algoritmos disponibles en la versión 5.2 de la herramienta ProM como marco de implementación de la Minería de Procesos.

2. CPN TOOLS COMO HERRAMIENTA PARA LA SIMULACIÓN DE PROCESOS DE NEGOCIO

CPN Tools se encuentra entre los simuladores basados en Redes de Petri que poseen la característica de incluir el lenguaje de simulación y el paquete de simulación. Estas herramientas combinan el entorno gráfico de diseño de los modelos de simulación y el lenguaje de programación, además de ofrecer capacidades de análisis gráficos y animación [13].

Según la bibliografía consultada [1], [3], [10], CPN Tools constituye una poderosa herramienta para modelar, editar, simular y analizar Redes de Petri Coloreadas (CPN por sus siglas en inglés) además de dar soporte al modelado de procesos simultáneos, donde los tokens tienen tipos asignados y las redes definidas pueden ser

extendidas con el tiempo, las jerarquías y manipuladas empleando el lenguaje de programación funcional subyacente, SML (Standard Markup Language). Otra importante funcionalidad que provee CPN Tools, es la habilidad de simular redes de una manera gradual controlada por el usuario, con el objetivo de evaluar caminos de ejecución específicos, y de manera automática, que es más rápida y permite generar información estadística.

Algunas de las ventajas de CPN Tools señaladas en [8], [1], [2], [3], [15], son las siguientes:

El usuario interactúa con CPN Tools sobre la base de la manipulación directa de la representación gráfica del modelo CPN, empleando técnicas de retroalimentación gráfica que mantienen al usuario informado del estado de la simulación, ya que se conoce exactamente la instancia que se ejecuta dentro del modelo.

La interacción de CPN Tools con las herramientas de Minería de Procesos permite obtener la información necesaria para el modelo de simulación a partir de los registros de eventos de los procesos automatizados con sistemas de información. De esta forma se obtiene el comportamiento espontáneo y real del proceso.

Obtener los datos para los modelos de simulación en CPN desde los registros de eventos, posibilita trabajar con la población de datos, ya que en estos registros se almacenan años de comportamiento del proceso, por tanto se evita el error de la muestra.

Como desventaja, según estos mismos autores, se podría mencionar el alto nivel de detalle requerido para modelar la manipulación de recursos y los aspectos relacionados con los tiempos de espera y las colas. Como consecuencia los modelos no pueden ser construidos fácilmente. El formalismo de Redes de Petri, aunque muy poderoso resulta difícil de comprender para usuarios inexpertos.

Es inevitable comentar que la modelación de entornos reales sea mediante la herramienta ARENA o CPN Tools, obliga a la combinación de las facilidades de los paquetes de simulación con el lenguaje de programación, para lograr el nivel de detalle en el modelo que permita recrear el comportamiento de procesos complejos.

2.1. Comparación cualitativa entre ARENA y CPN Tools

De manera general, en la bibliografía estudiada [8], [1], [9], se definen criterios de evaluación que permiten establecer comparaciones entre herramientas de simulación. A continuación en la tabla 1 se muestra un resumen comparativo de las herramientas ARENA y CPN Tools.

Tabla 1. Evaluación de herramientas para simulación de procesos de negocio.

CAPACIDAD DE MODELADO		
Criterios	ARENA	CPN Tools
Fácil construcción del modelo	+	+/-
Verificación de sintaxis	+/-	++
Nivel de detalle en su representación gráfica	+	++
Representación gráfica	+	-
CAPACIDAD DE SIMULACIÓN		
Animación	+	+
Análisis de escenarios	++	++
ANÁLISIS DE LOS RESULTADOS		
Estadísticas	+/-	++

A partir de los resultados presentados en la tabla 1 las herramientas se encuentran equilibradas teniendo en cuenta los aspectos utilizados para su comparación.

Destaca CPN Tools en la verificación de sintaxis pues esta se realiza a medida que se construye el modelo y no hay que esperar a terminar las modificaciones para conocer los errores de modelación introducidos por el analista.

El nivel de detalle que se muestra en CPN Tools favorece la modelación de procesos complejos, aunque dificulta el entendimiento para usuarios inexpertos. Se debe señalar que este alto nivel de detalle se puede alcanzar con la herramienta ARENA utilizando los más bajos niveles de modelado.

En el aspecto de fácil construcción del modelo, ARENA se destaca debido a la facilidad de construcción que brindan los niveles altos de modelado, pero su utilización está sujeta a entornos restringidos.

La representación gráfica en ARENA es más amigable, pero esta evaluación está inherentemente relacionada con los aspectos antes mencionados de nivel de detalle y el nivel de modelado utilizado. CPN Tools utiliza una representación gráfica discreta asociada a la condición de herramienta no propietaria.

En el aspecto de análisis de los resultados, CPN Tools destaca por la posibilidad de calcular para un mismo indicador diferentes estadísticas (cantidad de observaciones, valor promedio, desviación, varianza, valores mínimos y máximos, etc) dentro de un mismo experimento de simulación, teniendo en cuenta niveles de confianza de 90%, 95% y 99%. En ARENA, para cada indicador solo puede ser analizada una estadística a la vez.

3. COMPARACIÓN CUANTITATIVA ENTRE ARENA Y CPN TOOLS

Para el desarrollo del caso de estudio se toma como referencia el proceso *Diagnóstico de Pacientes*. El registro de eventos para este proceso es un ejemplo artificial basado en una clínica de atención a pacientes externos de la vida real de un hospital en Holanda. El proceso comienza con la primera visita del paciente a la clínica. Durante esta *Primera Visita* es tomada una muestra de sangre al paciente para ser analizada posteriormente en el laboratorio. En paralelo a la *Prueba de Laboratorio*, el paciente procede a realizarse un examen de *Rayos X*. La *Prueba de Laboratorio* y el examen de *Rayos X* pueden ocurrir en cualquier orden. Una vez realizada la prueba de laboratorio y los *Rayos X* el paciente es enviado a pedir una cita para un *Electrocardiograma* en caso de que lo necesite. Luego de estos exámenes, concluye la *Primera Visita* y el paciente tendrá que regresar para una visita de seguimiento en la clínica. Durante la *Segunda visita* el médico decide si es necesario realizar una *Resonancia magnética* o una *Tomografía computarizada*. Luego de la *Resonancia magnética* o la *Tomografía computarizada*, el proceso de diagnóstico concluye con una *Tercera Visita* a la clínica [11].

Obtención de los datos para la creación del modelo de simulación

Las técnicas tradicionales, como es el caso de los Estudios de Tiempo que utiliza ARENA, empleadas en la obtención de la información para la construcción del modelo de simulación, constituye uno de los métodos más costoso en tiempo y recursos, además de estar sujetos a la experticia de los analistas.

En cambio, el empleo de Minería de Procesos para la obtención de modelos de simulación, a partir de registros de eventos de procesos reales contenidos en sistemas de información, permite disminuir el tiempo invertido en la obtención de la información y los costos asociados a la contratación de asesores y analistas para la realización de Estudios de Tiempo. Permite además, eliminar los errores que se derivan de la percepción humana de la realidad, del trabajo con tamaños de muestra que no siempre reflejan la realidad del proceso y del empleo de probabilidades.

Utilizar la información almacenada en los registros de eventos de los sistemas de información para obtener modelos basados en la ejecución real del proceso, permite realizar la simulación sobre este modelo y no sobre una versión idealizada del proceso.

Construcción del modelo de simulación

La construcción de los modelos de simulación en ARENA está sujeta a la manera en que el analista percibe el proceso y requiere el empleo de probabilidades calculadas a partir de tamaños de muestras que no reflejan completamente la realidad.

CPN Tools se nutre de técnicas de Minería de Procesos que permiten obtener de manera automática los modelos de simulación a partir de registros de eventos inalterables en los cuales se pueden almacenar años de ejecución de un proceso de negocio, dando la posibilidad de realizar análisis con toda la población de los datos del proceso y aportando a los modelos representados una mayor semejanza respecto a la situación real.

Para la creación del modelo del proceso Diagnóstico de Pacientes en ARENA, se utilizó una combinación de los paneles Common, Basic Process, Blocks y Elements de ARENA. Se necesitó un Arrive, seis Server, un Branch, un Separate, un Batch, dos Decide, dos Tally, tres Assign, dos Delay, un Depart, un Simulate, un Statistics y un Tallies.

De manera general la construcción del modelo de simulación en ARENA, que se muestra en la figura 1, no se puede describir como sencilla, pues la complejidad del proceso y algunas de sus características hicieron necesario programar actividades a diferentes niveles, empleando varios paneles y un total de 18 bloques, lo cual dificulta el proceso de construcción, pues exige un conocimiento más amplio de la herramienta. En cambio, la obtención del modelo en Red de Petri para el proceso se clasifica como sencilla, pues el modelo, que se muestra en la figura 2, es obtenido de manera automática empleando técnicas de Minería de Procesos y posteriormente exportado a CPN Tools, lo cual, elimina el riesgo de introducir errores durante su modelado.

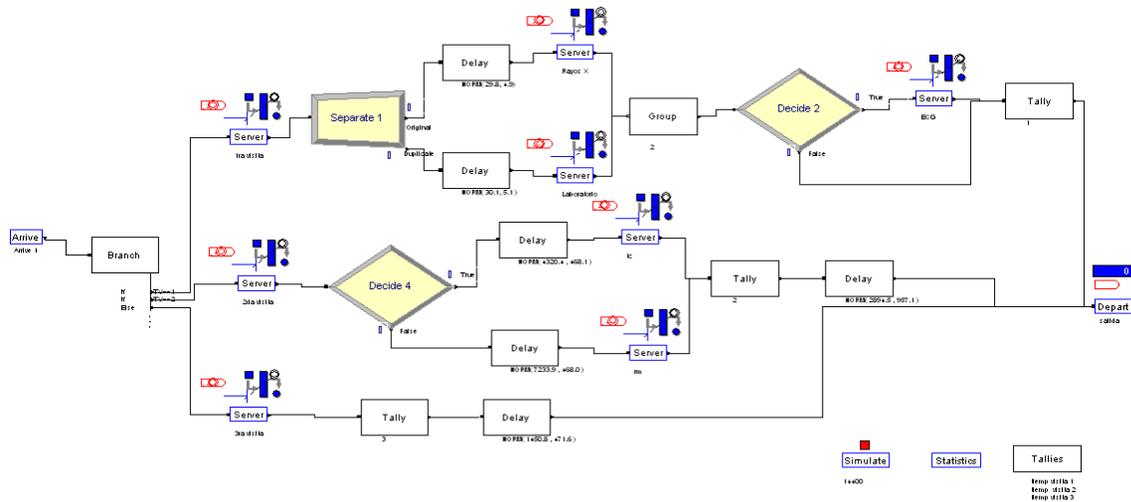


Figura 1. Representación del modelo de simulación en ARENA.

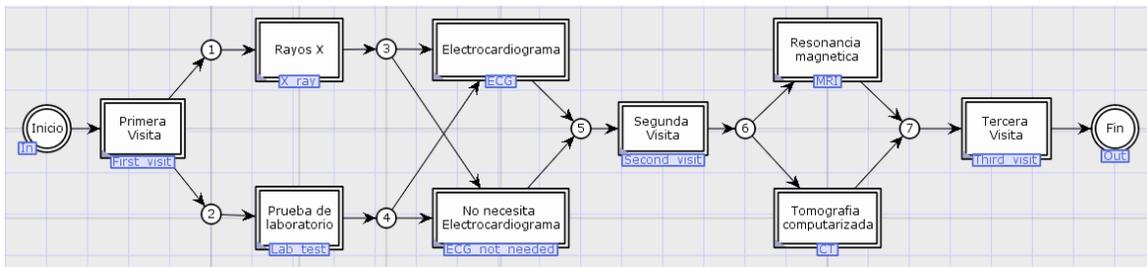


Figura 2. Representación del modelo de simulación en CPN Tools.

El empleo de técnicas informales de modelado, como es el caso de los diagramas de flujo que utiliza ARENA suele resultar de mayor facilidad y comprensión al no estar atadas a restricciones matemáticas, sin embargo, la principal desventaja es el no poder verificar los modelos obtenidos, con lo que es necesario esperar hasta tener el sistema desarrollado e implementado para verificar que todos sus módulos funcionan como se modelaron. En cambio, las técnicas de modelado formales, como es el caso de las Redes de Petri descansan sobre bases matemáticas que permiten que el modelo obtenido sea validado sintácticamente, permitiendo que el modelo sea analizado y verificado, asegurando así en mayor grado la ausencia de errores.

La verificación de sintaxis provee un descripción precisa e inequívoca del comportamiento del proceso modelado, por tanto, es necesario tenerla en cuenta por la importancia que tiene durante la modelación de un proceso, pues de alguna manera constituye una ayuda para el usuario, el cual necesita responder a preguntas tales como: ¿cuál fue el error? o ¿dónde se cometió el error? Para ello la mayoría de los paquetes de simulación cuentan con mecanismos que permiten establecer el cuál y el dónde y de su capacidad para detectar estos aspectos depende en gran medida la elección de uno u otro.

CPN Tools verifica la sintaxis de la red de manera automática durante su construcción o al ser cargada automáticamente. Cuando se realizan cambios en la red, el chequeo de sintaxis verifica solamente lo partes de la red que son afectadas por el cambio, por ejemplo, cuando una declaración es cambiada, el verificador de sintaxis no vuelve a chequear todas las declaraciones, sino que verifica solamente las declaraciones y las inscripciones de red que dependen de la declaración que ha sido modificada. Esto permite que el usuario haga los ciclos pequeños de la edición y la simulación sin tener que esperar hasta tener el sistema desarrollado.

Análisis del proceso con la simulación

CPN Tools posee un mecanismo denominado *Retroalimentación de simulación*, que destaca como una de sus principales ventajas y es que da la posibilidad de conocer exactamente, a través del marcado de la red, que instancia(s) se encuentra(n) en el proceso en un momento determinado y seguir su comportamiento, conocer

el momento exacto de inicio y fin de las actividad para cada instancia y personificar los recursos, es decir, conocer con exactitud el/los recurso(s) que están siendo utilizados en un momento dado.

Al igual que CPN Tools, ARENA permite seguir el recorrido de las entidades durante su paso por el proceso y conocer su comportamiento, la desventaja, es no poder diferenciar entre ellas y conocer los momentos exactos, por ejemplo, de su llegada y/o el inicio de una actividad. Lo mismo sucede con los recursos y se debe a que en ARENA no es posible personificarlos, o sea, ARENA permite definir un grupo de recursos para una actividad, pero no se conoce cuál de ellos la ejecuta en un momento determinado.

3.1. Análisis y validación de los resultados de la simulación

La aplicación de técnicas de Minería de Procesos en ProM permite obtener, a partir del registro de eventos de un proceso automatizado, el comportamiento real de todos los indicadores de rendimiento asociados a los tiempos y la utilización de los recursos, pero no es una herramienta con fines de simulación, por lo cual, no permite variar parámetros dentro del modelo en función de un rediseño. Por ello, es necesario auxiliarse en las facilidades de simulación que brinda la herramienta CPN Tools, la cual, se nutre de técnicas de Minería de Procesos para la obtención de los modelos de simulación, permitiendo importar los modelos descubiertos automáticamente desde ProM para posteriores análisis².

En la tabla 2 se muestra el comportamiento de los indicadores de rendimiento obtenidos con la simulación en ARENA. Se incluyen además los resultados obtenidos con Minería de Procesos y con la simulación en CPN Tools. En este caso, con fines comparativos, los valores de los indicadores de rendimiento *Tiempo de permanencia en el sistema (TPS)* y *Tiempos de espera* son obtenidos con Minería de Procesos, los valores correspondientes al indicador *Utilización de los recursos* se obtienen a través de la simulación del proceso en CPN Tools, se incluye en ambos casos el valor correspondiente al indicador *Tiempo total de procesamiento*. Estos resultados son analizados y comparados con los datos reales almacenados en el registro de eventos.

Tabla 2. Comportamiento de los indicadores de rendimiento.

Variable respuesta		UM	Registro de eventos	ProM	CPN Tools	ARENA	
Cantidad de arribos		uf	235	235	242	242	
TPS primera visita		min	127.42	127.42		120.32	
TPS segunda visita		min	5871.13	5871.13		5608.4	
TPS tercera visita		min	29.64	29.64		30.86	
Tiempo total de procesamiento		min	10284.43	10284.43	10327.31	11374.0	
Tiempos de espera	Primera visita	min	0	0		0.44	
	Rayos X	min	29.49	29.49		0	
	Examen de laboratorio	min	29.77	29.77		0	
	ECG	min	0.02	0.02		0	
	Segunda visita	min	2937.24	2937.24		0.02	
	RM	min	7296.72	7296.72		0	
	TC	min	4280.11	4280.12		0	
	Tercera visita	min	1372.65	1372.65		0.58	
Utilización de recursos	Ginecólogos	Primera visita	%			76.01	
		Segunda visita	%			39.63	
		Tercera visita	%			14.67	
	Cardiólogos	ECG	%			24.25	12.26
	Radiólogos	Rayos X	%			33.60	27.61
		RM	%			13.41	
		TC	%			19.48	
	Técnicos	Examen de lab.	%			33.51	11.27

² CPN Tools permite la obtención de los modelos de simulación de manera automática utilizando Minería de Procesos y de manera manual a través del modelado del proceso.

Como se observa en la tabla 2, los datos almacenados en el registro de eventos, se corresponden con los valores obtenidos en ProM luego de aplicar las técnicas de Minería de Procesos, lo que significa que el modelo de simulación obtenido con Minería de Procesos es capaz de reflejar la realidad del proceso registrada en el log.

Los tiempos de espera obtenidos con ARENA son casi nulos y difieren significativamente de los almacenados en el registro de eventos. Esto está relacionado con la manera en que el analista concibe el modelo, por lo cual, no siempre es posible reflejar la realidad contenida en el registro de eventos.

Análisis de normalidad.

Por su importancia para la realización de análisis posteriores se estudia la normalidad de los valores reales almacenados en el registro de eventos, a partir de los cuales se elaboraron los modelos de simulación en ARENA y en CPN Tools. Con esta finalidad se aplica la prueba Kolmogorov-Smirnov para bondad de ajuste. Se trabaja para un nivel de confianza del 95% y un nivel de significancia (α) de 0.05.

Partiendo de la necesaria aleatoriedad de los datos se formula la siguiente hipótesis.

Hipótesis:

- ❖ **H₀**: Los datos ajustan a una distribución normal.
- ❖ **H₁**: Los datos no ajustan a la distribución normal.

En la tabla 3 se muestran los resultados de la prueba obtenidos para los valores correspondientes a los indicadores *Tiempo total de procesamiento* y *Tiempo de permanencia en el sistema*.

Tabla 3. Resultados obtenidos para la prueba Kolmogorov-Smirnov de bondad de ajuste.

VARIABLE	N	MEDIA	DESV.EST.	VALOR P
Tiempo total de procesamiento	235	10284.43	1291.98	0.749
TPS primera visita	235	127.42	15.29	0.074
TPS segunda visita	235	5871.13	880.75	0.199
TPS tercera visita	235	29.64	4.14	0.229

Como se observa en la tabla 3, los valores P obtenidos para todos los indicadores son superiores al nivel de significancia preestablecido, por tanto, se acepta la hipótesis H_0 y se concluye que los valores analizados se ajustan a una distribución normal.

Comparación de medias

Para determinar si los resultados obtenidos con CPN Tools y ARENA, poseen diferencias significativas con respecto a la información almacenada en el registro de eventos, se utiliza prueba de hipótesis con ayuda del software Minitab. Para la comparación se emplean los resultados correspondientes a los indicadores *Tiempo total de procesamiento*, asociado al tiempo promedio que tarda un paciente en ejecutar el proceso de diagnóstico y *Tiempo de permanencia en el sistema*, asociado al tiempo promedio que permanece un paciente en las actividades *Primera Visita*, *Segunda Visita* y *Tercera Visita*. Se trabaja para un nivel de confianza del 95% y un nivel de significancia (α) de 0.05.

Primeramente se plantean las hipótesis estadísticas para verificar diferencias significativas entre los resultados obtenidos con CPN Tools y la información real que se encuentra en el registro de eventos del proceso.

Hipótesis:

- ❖ **H₀**: $\mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$. No existe diferencia significativa entre las medias.
- ❖ **H₁**: $\mu_1 \neq \mu_2$ $\mu_1 - \mu_2 \neq 0$. Existe diferencia significativa entre las medias.

En la tabla 4 se muestran los resultados obtenidos para la prueba T de comparación de dos medias, aplicada al valor obtenido con la simulación del proceso en CPN Tools y al dato real almacenado en el registro de eventos para el indicador *Tiempo total de procesamiento*.

Tabla 4. Resultados obtenidos para la prueba T de comparación de dos medias. CPN Tools/Registro de eventos.

MUESTRA	TAMAÑO DE MUESTRA	MEDIA	ESTIMADO DE LA DIFERENCIA	VALOR P
Registro de eventos	235	10284.43	-43	0.798
CPN Tools	235	10327.31		

De la tabla 4 se puede concluir que la diferencia existente entre el valor medio obtenido para el indicador *Tiempo total de procesamiento*, utilizando simulación en CPN Tools y el valor real almacenado en el registro de eventos, no se considera estadísticamente significativa, dado que el valor P para la prueba es superior al nivel de significancia establecido y por tanto se acepta la hipótesis nula (H_0).

De igual manera se procede con el valor obtenido de la simulación del proceso en ARENA para el indicador *Tiempo total de procesamiento*. Las hipótesis estadísticas para verificar diferencias significativas entre los resultados obtenidos con ARENA y la información real que se encuentra en el registro de eventos del proceso son las siguientes:

Hipótesis:

- ❖ $H_0: \mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$. No existe diferencia significativa entre las medias.
- ❖ $H_1: \mu_1 \neq \mu_2$ $\mu_1 - \mu_2 \neq 0$. Existe diferencia significativa entre las medias.

En la tabla 5 se muestran los resultados obtenidos para la prueba T de comparación de dos medias, aplicada al valor obtenido con la simulación del proceso en ARENA y al dato real almacenado en el registro de eventos.

Tabla 5. Resultados obtenidos para la prueba T de comparación de dos medias. ARENA/Registro de eventos.

MUESTRA	TAMAÑO DE MUESTRA	MEDIA	ESTIMADO DE LA DIFERENCIA	VALOR P
Registro de eventos	235	10284.43	-1090	0.000
ARENA	235	11374		

De la tabla 5 se puede concluir que la diferencia que existe entre el resultado obtenido con la simulación en ARENA y el almacenado en el registro de eventos se considera estadísticamente significativa, pues el valor P obtenido para la prueba es inferior al nivel de significancia establecido y por tanto se acepta la hipótesis H_1 .

En las tablas 6 y 7 se muestran los valores correspondientes al indicador *Tiempo de permanencia en el sistema (TPS)* para las actividades *Primera Visita*, *Segunda Visita*, *Tercera Visita*, y el valor P obtenido al aplicar la prueba T de comparación de dos medias a los resultados obtenidos con Minería de Procesos en ProM y con simulación en ARENA. El planteamiento de las hipótesis es el siguiente:

Hipótesis:

- ❖ $H_0: \mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$. No existe diferencia significativa entre las medias.
- ❖ $H_1: \mu_1 \neq \mu_2$ $\mu_1 - \mu_2 \neq 0$. Existe diferencia significativa entre las medias.

Tabla 6. Resultados obtenidos para la prueba T de comparación de dos medias. ARENA/Registro de eventos.

VARIABLE RESPUESTA	REGISTRO DE EVENTOS	ARENA	VALOR P
TPS primera visita	127.42	120.32	0
TPS segunda visita	5871.13	5608.4	0
TPS tercera visita	29.64	30.86	0

Tabla 7. Resultados obtenidos para la prueba T de comparación de dos medias. ProM/Registro de eventos.

VARIABLE RESPUESTA	REGISTRO DE EVENTOS	PROM	VALOR P
TPS primera visita	127.42	127.42	1
TPS segunda visita	5871.13	5871.13	1
TPS tercera visita	29.64	29.64	1

Como se observa en la tabla 6, los valores P, resultado de la comparación de los tiempos medios de permanencia de los pacientes en las actividades *Primera Visita*, *Segunda Visita* y *Tercera Visita*, obtenidos

con la simulación en ARENA y los valores reales almacenados en el registro de eventos, es igual a 0, lo que significa que estadísticamente la diferencia entre estos resultados se considera significativa.

En cambio, en la tabla 7 se observa que los valores obtenidos con Minería de Procesos se corresponden totalmente con los datos reales almacenados en el registro de eventos. Este último resultado demuestra las ventajas de emplear técnicas de Minería de Procesos y simulación sobre Redes de Petri en CPN Tools, para la obtención y el análisis de modelos de simulación de procesos reales que se encuentren automatizados. Estas técnicas permiten disminuir el tiempo y los costos asociados a la obtención de la información, obtener modelos más cercanos a la realidad y resultados más fiables, pues disminuyen el riesgo de introducir errores al modelo y los análisis se realizan sobre la población de datos almacenados en los registros de eventos.

4. CONCLUSIONES

Se identifican las Redes de Petri como una poderosa herramienta para el modelado de procesos de negocio, destacándose de otros métodos por su semántica, su representación gráfica, por la expresividad de información que contienen sus estructuras, así como por su fundamento matemático y las técnicas de análisis que ofrecen, entre ellas, la simulación. Es un lenguaje de modelación lo suficientemente versátil para ser utilizado en la práctica por sistemas de gran tamaño y complejidad.

El empleo de técnicas de Minería de Procesos y simulación en CPN Tools, para la obtención y el análisis de modelos de simulación de procesos reales, disminuye el riesgo de introducir errores al modelo, permite realizar los análisis sobre la población de los datos almacenados en el registro de eventos y disminuir el tiempo y los costos asociados a la obtención de la información.

Con la validación de los resultados quedaron demostradas las ventajas de emplear técnicas de Minería de Procesos y simulación en CPN Tools, sobre las técnicas tradicionales, para la obtención y el análisis de modelos de simulación de procesos reales previamente automatizados, permitiendo obtener modelos más cercanos a la realidad, resultados más fiables y potenciar los análisis de rendimiento en tiempo real.

RECEIVED NOVEMBER, 2014

REVISED MARCH, 2015

REFERENCIAS

- [1] AILENEI, I.-M. (2011): Process Mining Tools: A Comparative Analysis. **Partial fulfillment of the requirements for the degree of Master of Science in Business Information Systems**, Eindhoven University of Technology, Netherlands.
- [2] ADRIANSYAH, A. et al. (2012): Mining Process Performance from Event Logs. En: **Business Process Management Workshops**, 217-218. Eindhoven University of Technology. Springer-Verlag, Berlin.
- [3] BOSE, J. C. (2012): **Process Mining in the large. Preprocessing, Discovery and Diagnostics**. Eindhoven University of Technology, Netherlands.
- [4] IGNA, G. (2013): Performance Analysis of Real-Time Task Systems using Timed Automata. **Trabajo presentado en opción al grado académico de Doctor en Ciencias**. Institute for Programming research and Algorithmic, Netherlands.
- [5] INTERNATIONAL STANDARD ISO/IEC 15909-1 (2004) **Software and system engineering. High-level Petri nets, Part 1: Concepts, definitions and graphical notation**. International Organization for Standardization, Ginebra, Suiza.
- [6] INTERNATIONAL STANDARD ISO/IEC 15909-2. (2011): **Software and Systems Engineering – High-level Petri Nets, Part 2**. International Organization for Standardization, Ginebra, Suiza.
- [7] INTERNATIONAL STANDARD ISO/IEC 15909-3. (2014): **Software and Systems Engineering – High-level Petri Nets, Part 3: Petri Nets Extensions**. International Organization for Standardization, Ginebra, Suiza.
- [8] JENSEN et al. (2007): **Coloured Petri Nets and CPN Tools for modeling and validation of concurrent systems**. Springer-Verlag, Berlin.
- [9] MANS, R. (2012): **Digitalizing the Dental Oral Implant Treatment Workflow: Quantification of the Impact on Precision and Efficiency**. Eindhoven University of Technology, Netherlands.
- [10] MAYORGA, S. A. (2013): Mejoramiento de procesos con técnicas de Minería de Procesos, simulación y optimización. Caso de estudio. **NOVÁTICA, Revista de la Asociación de Técnicos de Informática**, 25, 270 – 277.

- [11] ROZINAT, R. S. M., M. SONG, AND W.M.P. VAN DER AALST (2009): **Discovering Simulation Models**. Eindhoven University of Technology, Netherlands.
- [12] VAN DER AALST, W. M. P. et al. (2010): Business Process Simulation. En: Silver.B. **Handbook on Business Process Management**, 313-339. Springer-Verlag, Berlín.
- [13]VAN DER AALST, W. M. P. et al. (2011): Manifiesto sobre Minería de Procesos. IEEE Task Force on Process Mining. Available in: www.win.tue.nl/ieeetfpm, www.springer.com.
- [14]VAN DER AALST, W. M. P. and BOSE, R. P. J. C., (2012): Process Diagnostics using Trace Alignment: Opportunities, Issues, and Challenges. **Information Systems**, 37(2), 117-141.
- [15]VAN DER AALST, W.M.P et al (2013) Business Process Simulation Survival Guide. En: Silver.B. **Handbook on Business Process Management**, Part II, 337-370. Springer-Verlag, Berlin.