

INCEPTION: Gestión inteligente del cambio en minería de procesos sustentada por la nube*

Víctor Gallego Fontenla¹[0000-0002-4149-919X], Jorge García de Quirós²,
Alberto Bugarín¹[0000-0003-3574-3843], Pedro Álvarez²[0000-0002-6584-7259], Juan C. Vidal¹[0000-
0002-8682-6772], Javier Fabra²[0000-0001-5549-7649], Manuel Lama¹[0000-0001-7195-6155]

¹ Centro Singular de Investigación en Tecnoloxías Intelixentes
Universidade de Santiago de Compostela, 15782 Santiago de Compostela, España
{victorjose.gallego, alberto.bugarin.diz, juan.vidal, manuel.lama}@usc.es

² Departamento de Informática e Ingeniería de Sistemas
Universidad de Zaragoza, 50009 Zaragoza, España
{jgarciaqg, alvaper, jfabra}@unizar.es

Resumen. En este artículo se presenta el proyecto INCEPTION, cuyo objetivo es la transferencia de la tecnología de detección y descripción del cambio de los modelos de proceso, desarrollada en un proyecto previo de I+D+i, para que opere de forma eficiente en entornos masivos de datos para procesos reales, complejos y no estructurados. En el artículo se describirá el contexto en el que nace el proyecto, así como los objetivos y las actividades que se realizarán para llevar a cabo la transferencia a las plataformas comerciales de minería de procesos, en especial, a la plataforma InVerbis Analytics.

Keywords: Detección del cambio; Minería de procesos; Nube.

1 Introducción

Las técnicas de minería de procesos han surgido como una forma de entender el comportamiento de una organización, extrayendo el conocimiento a partir del análisis de los registros de un sistema de, información, compuestos por trazas de eventos, y con el fin de descubrir, monitorizar y enriquecer los procesos reales [1]. Entre las técnicas de minería de procesos con una mayor implantación en las plataformas comerciales, destacan las siguientes: (i) *descubrimiento de procesos*, cuyo objetivo es obtener automáticamente un modelo de proceso que sea capaz de reproducir las trazas del registro de eventos (completitud) representando de la forma más fiable y concisa (precisión y simplicidad) las relaciones entre actividades. Todas las plataformas comerciales usan algoritmos de descubrimiento de procesos [2], puesto que las analíticas extraídas a partir de los modelos constituyen la base en la que se fundamenta la minería de procesos; (ii) *simplificación del modelo de procesos*, cuyo objetivo es modificar los modelos de los procesos para hacerlos más comprensibles, sobre todo para procesos no estructurados

* Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación a través de los proyectos PDC2021-121072-C21 y PDC2021-121072-C21.



que requieren de modelos con un alto grado de completitud y en los que los algoritmos de descubrimiento obtienen modelos muy complejos. Aunque en los últimos años han aparecido propuestas que integran el comportamiento infrecuente en actividades abstractas que ayudan a ofrecer una visión completa del modelo [3], la mayoría de las plataformas comerciales usan una estrategia de visualización incremental del modelo en la que el comportamiento observado se va incorporando al modelo en función de su frecuencia; (iii) *comprobación de la conformidad del proceso*, que tiene como objetivo describir y cuantificar las diferencias entre el modelo del proceso y el comportamiento observado en el registro de eventos. Las plataformas comerciales utilizan estrategias subóptimas a la hora de buscar el alineamiento entre las trazas y el modelo de proceso, ya que son suficientes para orientar qué tipo de trazas se desvían del modelo y en qué medida tiene lugar este desvío; y (iv) *visualización y descripción de las analíticas del proceso/modelo (enriquecimiento)*, incluyendo descripciones textuales cuyo objetivo es complementar la información que proporcionan las utilidades de visualización de modelos y analíticas de proceso para facilitar la navegación del usuario a través del modelo y permitir centrar su atención en lo relevante del proceso [4]. Ninguna plataforma comercial actual incorpora componentes para la generación de descripciones textuales.

Estas técnicas tienen como entrada ((ii) a (iv)) o generan ((iv)) un modelo de proceso que consideran estático, es decir, existe un único modelo que da respuesta a todo el comportamiento del registro de eventos. Sin embargo, la naturaleza de los procesos es dinámica, sobre todo en procesos no estructurados, en los cuales el modelo consiste en la combinación de varios modelos del proceso a lo largo del tiempo. Por lo tanto, identificar y describir cuándo el modelo de un proceso ya no representa el comportamiento observado, porque el modelo ha cambiado de forma súbita o gradual [5] es uno de los principales aspectos a considerar en el análisis de un proceso. Sin embargo, ninguna plataforma comercial incorpora técnicas de detección y descripción de cambio (*concept drift*, en inglés) en los modelos, lo cual penaliza las conclusiones y resultados extraídos por las otras técnicas de minería de procesos.

En este artículo se describen los objetivos y las actividades del proyecto de Prueba de Concepto “*Gestión inteligente del cambio en minería de procesos sustentada por la nube*”, que adapta los algoritmos de detección y descripción del cambio en modelos de proceso, desarrollados en un proyecto previo de I+D+i, a procesos reales, complejos y no estructurados y las despliega en la nube para que sean ejecutadas de forma eficiente en entornos de datos masivos y se puedan integrar de modo flexible con las plataformas comerciales de minería de procesos.

2 Objetivos del proyecto

El objetivo principal del proyecto consiste en la adaptación de la tecnología de gestión del cambio en modelos de procesos a las necesidades de los entornos de negocio reales con grandes volúmenes de datos, facilitando la detección y descripción textual de dichos cambios. Además de este objetivo general se han identificado los siguientes objetivos específicos:

- *Adaptación de los algoritmos de detección del cambio en modelos de proceso a los resultados de la experimentación con registros de procesos de clientes reales*, en los que es necesario tratar con la complejidad de los modelos del proceso y la variabilidad de las trazas. Esta adaptación consistirá en integrar varios algoritmos de descubrimiento de procesos, medidas eficientes de completitud y de precisión, y algoritmos de agrupamiento de trazas para comprobar cuáles de ellos se ajustan mejor a las necesidades de los clientes reales.
- *Adaptación de la tecnología de process-to-text para la descripción en lenguaje natural de los cambios en el modelo que representa la conducta observada de un proceso*. Estas descripciones se centrarán (i) en las diferencias estructurales entre el modelo previo y el nuevo modelo; y (ii) en las diferencias entre los indicadores de proceso asociados a las actividades y trazas, adaptando los componentes con los que se describen las analíticas de proceso.
- *Diseño y desarrollo de una infraestructura orientada a servicios que integre los algoritmos para la detección y descripción del cambio de los procesos*. Esta infraestructura ofrecerá su funcionalidad conforme a la aproximación orientada a servicios y estará desplegada sobre recursos de cómputo y almacenamiento de múltiples proveedores de nube. Además, dispondrá de mecanismos que faciliten el proceso de migración de sus componentes entre proveedores en base a criterios de rendimiento y/o costes. El objetivo es disponer de una solución final sea flexible y fácilmente adaptable a distintos entornos tecnológicos y a las particularidades técnicas de estos, siempre desde la perspectiva de la nube.
- *Licenciamiento y transferencia de la propuesta orientada sobre todo a empresas que ya disponen de tecnología de minería de procesos*. Desde un punto de vista tecnológico la propuesta de prueba de concepto es un producto que no depende de terceros ni de ninguna tecnología adicional y que podrá ser implantado en empresas con dicha necesidad. No obstante, inicialmente se considera como primera opción un acuerdo con InVerbis Analytics, empresa de base tecnológica de la Universidad de Santiago de Compostela que dispone de una plataforma comercial de minería de procesos.

3 Tareas a desarrollar en el proyecto

Las tareas a realizar en el proceso para alcanzar estos objetivos son las siguientes:

- Desarrollar las adaptaciones necesarias para que el algoritmo de detección del cambio sea capaz de tratar con grandes volúmenes de datos y con procesos no estructurados, lo que incluye la integración de algoritmos de agrupamiento de trazas para reducir la complejidad del problema. Además, con el fin de facilitar al usuario la comprensión de los cambios que han ocurrido y el impacto que tienen sobre el proceso, se realizará la adaptación de la tecnología *process-to-text* para describir textualmente los cambios que tengan lugar, así como el desarrollo de una interfaz gráfica para visualizar esos cambios en el tiempo con sus descripciones.
- Realizar un trabajo de consultoría para analizar y evaluar las tecnologías y servicios de nube que podrían aplicarse en el desarrollo de la infraestructura de servicios. Esta consultoría tecnológica incluye también un estudio de la nube desde una perspectiva

económica y un análisis completo sobre el contexto tecnológico en el que se enmarcan las empresas a las que podría estar destinado el producto final. Posteriormente, y a partir de los resultados de la consultoría, se realizará una adaptación arquitectural y desarrollo de la infraestructura de servicios de minería de procesos en la que se integrarán los algoritmos (adaptados) de detección y descripción de cambios en modelos de proceso, así como otros algoritmos necesarios para su correcto funcionamiento. La infraestructura constará de un conjunto de servicios que se desplegarán sobre múltiples proveedores de nube, en función de criterios de rendimiento y coste, y que serán fácilmente integrables en cualquier entorno empresarial. Además, se programarán mecanismos que faciliten y automaticen, en la medida de lo posible, el aprovisionamiento dinámico de los recursos necesarios para ejecutar los algoritmos y gestionar y almacenar los datos involucrados.

- Realizar una validación comercial de la propuesta para demostrar la viabilidad e interés de la tecnología y adaptar y/o configurar sus módulos de detección y descripción del cambio en procesos y de procesamiento en la nube. Se comenzará por realizar una segmentación e identificación de los posibles clientes; en este contexto se entiende que los principales clientes de la propuesta son las empresas que ofrecen consultoría sobre minería de procesos, concretamente, las que poseen herramientas comerciales de minería de procesos. Una vez hayan sido identificados y analizados los posibles clientes y la tipología de los procesos, se seleccionarán cuatro pilotos que se desarrollarán a lo largo de un semestre con el fin de detectar las posibles mejoras que es necesario introducir en la propuesta para adaptarla a las necesidades de los clientes reales.
- Realizar la transferencia efectiva de la tecnología de la propuesta de prueba de concepto, incluyendo desde el modelo de explotación y monetización hasta la negociación de los posibles acuerdos con empresas que disponen y/o utilizan plataformas de minería de procesos, pasando por el licenciamiento y registro de la tecnología.

References

1. W.M.P. van der Aalst (2011): *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. Springer.
2. W. van der Aalst (2016): *Process mining software*. *Process Mining*, pp. 325-352.
3. D. Chapela-Campa, M. Mucientes, M. Lama (2020): *Understanding complex process models by abstracting infrequent behavior*. *Future Gener. Comput. Syst.* 113:428–440.
4. Y. Fontela-Seco, M. Lama, A. Bugarín (2021): *Process-To-Text: A framework for the quantitative description of processes in natural language*. *Trustworthy AI - Integrating Learning, Optimization and Reasoning*, pp. 212—129.
5. V. Gallego-Fontenla, J.C. Vidal, M. Lama (2019): *A conformance checking-based approach for drift detection in business processes*. *IEEE Transactions on Services Computing* 16(1):13—26.