

Tipo de artículo: Artículo original
Temática: Técnicas de programación
Recibido: 15/01/2017 | Aceptado: 22/01/2018

Algoritmo para inferir Redes de Interacción Social a partir de la ejecución de un proceso

Algorithm to infer Social Interaction Networks from the execution of a process

Leansi Vega Rouco*, Anel Wilfredo Fernández Rodríguez, Damián Pérez Alfonso

Facultad 3. Universidad de las Ciencias Informáticas, Carretera a San Antonio de los Baños, Km 2.5, Torrens, Boyeros, La Habana, Cuba. CP.: 19370

*Autor para correspondencia: leansi@uci.cu

Resumen

La minería de procesos es una disciplina de investigación que proporciona técnicas para descubrir, monitorear y mejorar los procesos de negocio en las organizaciones. En la minería de procesos, la inferencia de redes de interacción social para un proceso global es un elemento importante en el análisis del comportamiento de las interacciones de los recursos, pero cuando dichas interacciones ocurren en procesos complejos se dificulta la comprensión de los mismos en su dimensión social. Este análisis se enriquece cuando se realiza a nivel de subprocesos. El presente trabajo describe un algoritmo de inferencia de redes de interacción social a partir de eventos con alto número de interacciones entre los recursos, que contribuye a facilitar la comprensión de procesos complejos en su dimensión social. Los principales aportes de esta investigación son: la inferencia de redes de interacción social a nivel de subprocesos, el cálculo de los factores popularidad, eficiencia y sobrecarga; un complemento para la herramienta ProM que realiza el algoritmo diseñado. El cálculo de las métricas conocidas para inferir redes de interacción social y las medidas de centralidad permitieron la inferencia de redes de interacción social con el propósito de determinar los factores popularidad, eficiencia y sobrecarga de un recurso en la red. La correctitud del algoritmo propuesto se mostró mediante la realización de un caso de estudio, obteniendo resultados acordes al objetivo de la investigación y relacionados con sus principales aportes.

Palabras clave: Análisis de redes sociales, comprensión de procesos, métricas de inferencia, minería de procesos.

Abstract

Process mining is a research discipline that provides techniques for discovering, monitoring and improving business processes in organizations. In process mining, the inference of social interaction networks for a global process is an important element in the behavior analysis of resource interactions, but when such interactions occur in complex processes it is difficult to understand them in their social dimension. This analysis is enriched when performed at the subprocesses. The present work describes an algorithm of social interaction networks inference from events with high number of interactions between resources, which contributes to facilitate the understanding of complex processes in their social dimension. The main contributions of this research are: social interaction networks inference at subprocess level, calculation of factors popularity, efficiency and overload; An add-on for the ProM tool that performs the designed algorithm. The calculation of known metrics to infer social interaction networks and measures of centrality allowed the inference of social interaction networks with the purpose of determining the factors popularity, efficiency and overload of a resource in the network. The correctness of the proposed algorithm was shown by the realization of a study case, obtaining results according to the objective of the research and related to its main contributions.

Keywords: *Inference metrics, process mining, processes comprehension, social network analysis.*

Introducción

La inferencia de redes de interacción social constituye un problema de marcado interés en la comunidad científica durante los últimos años. Estudiar el modo en que las personas construyen sus relaciones y cómo estas interacciones desarrollan diversas potencialidades y patrones se le denomina Análisis de Redes Sociales (ARS) y puede contribuir a explicar la ocurrencia de ciertos fenómenos sociales. El entorno empresarial es uno de los contextos donde las personas interactúan a través de los procesos que ejecutan. La dimensión social de un proceso abarca estas interacciones, las cuales constituyen las relaciones sociales que se pueden establecer entre los recursos involucrados (Pérez, 2006).

Las empresas de hoy en día optan por gestionar sus procesos mediante sistemas informáticos. Las ejecuciones de estos procesos se registran en dichos sistemas en trazas de eventos. El hecho de que a partir de la información contenida en las trazas se pueda identificar el proceso, se le denomina Minería de Procesos (MP) o minería de flujos de trabajo (Agrawal, Gunopulos y Leymann, 1998).

La MP es una disciplina de investigación, que abarca un conjunto de técnicas para el descubrimiento automático de procesos, la verificación de conformidad, la construcción automática de modelos de simulación, la extensión de modelos, la reparación de modelos, la predicción de casos, las recomendaciones basadas en historia y la minería de redes sociales (Van der Aalst, Adriansyah y van Dongen, 2012).

Social Network Miner (SNM)¹ o Minería de Redes Sociales por sus siglas en inglés, es un complemento para la herramienta con capacidad de MP ProM que permite inferir redes de interacción social aplicando técnicas de MP. Este complemento permite realizar el ARS para un proceso global, siendo engorroso analizar una red con alto número de interacciones entre los recursos. Algunos factores que se ven afectados en el análisis de un proceso global están relacionados con el conocimiento de los recursos con mayor probabilidad de extender el proceso, aquellos que son indispensables en el proceso, los que presentan sobrecarga y los recursos y flujos de información que puedan ser optimizados. A partir de todo lo antes señalado se puede afirmar que el alto número de interacciones entre los recursos que ocurren en procesos complejos dificultan la comprensión de los mismos en su dimensión social. La inferencia de redes de interacción social a partir de registros de eventos constituye un criterio que aporta información relevante para la comprensión de procesos complejos (Pérez, 2015).

En esta investigación se presenta un algoritmo que combina un conjunto de técnicas de MP para inferir redes de interacción social a partir de eventos con alto número de interacciones entre los recursos, que contribuye a facilitar la comprensión de procesos complejos en su dimensión social.

Materiales y métodos

El algoritmo desarrollado comprende dentro de su base conceptual los términos de minería de procesos, teoría de grafos y las medidas centralidad, cercanía e intermediación para su correcto funcionamiento. Mediante la MP el algoritmo permite la inferencia de redes de interacción social mientras que el ARS puede realizarse haciendo uso de la teoría de grafos y las medidas centralidad, cercanía e intermediación.

La entrada del algoritmo es un registro de eventos en formato XES² que debe ser coherente con la definición 1 y 2 (tomadas de (Aalst, 2011)). Este registro de eventos representa un subproceso del proceso global y está compuesto por uno o varios eventos dentro de cada traza, debiendo ser coherente también con la definición 3 (tomada de (Van

¹ Social Network Miner: <http://www.processmining.org/online/snminer>

² eXtensible Event Stream: <http://www.xes-standard.org/>

der Aalst, Weijters y Maruster, 2004)). Es importante para la aplicación del algoritmo contar con un registro de eventos lo más completo posible y con todos los valores de sus atributos.

DEFINICIÓN 1 Eventos: *Sea E el universo de eventos. Los eventos pueden caracterizarse por varios atributos. Siendo AN el conjunto de nombres de atributos, para cualquier evento $e \in E$ y un nombre $n \in AN$: $\#n(e)$ es el valor del atributo n para el evento e . E^* representa el conjunto de todas las secuencias finitas sobre E . Una secuencia finita sobre E de longitud n es un mapeo $\sigma \in 1, \dots, n \rightarrow E$.*

DEFINICIÓN 2 Traza, registro de eventos: *Sea C el universo de casos. Para cualquier caso $c \in C$ y un nombre $n \in AN$: $\#_n(c)$ es el valor del atributo n para el caso c . Cada caso tiene un atributo obligatorio $\text{trace}\# \text{trace}(c) \in E^*$. $\hat{c} = \# \text{trace}(c)$ es un atajo para referirse a la traza de un caso.*

- Una traza es una secuencia finita de eventos $\sigma \in E^*$ tal que cada evento aparece una sola vez, o sea, para $1 \leq i \leq j \leq |\sigma|$: $\sigma(i) \neq \sigma(j)$.
- Un registro de eventos es un conjunto de casos $L \subseteq C$ tal que cada evento aparece al menos una vez. Si el registro de eventos contiene marcas de tiempo, los eventos deben ser ordenados dentro de una traza según esas marcas: $\forall c \in L, i \text{ y } j \text{ tal que } 1 \leq i < j \leq |\hat{c}|$: $\# \text{time}(\hat{c}(i)) \leq \# \text{time}(\hat{c}(j))$.

DEFINICIÓN 3 Subproceso: *Es una encapsulación de las actividades del negocio que representa una unidad de trabajo lógica y coherente. Los subprocesos tienen sus propios atributos y metas, pero contribuyen a alcanzar la meta general del proceso. Un subproceso es también un proceso y la mínima expresión de un subproceso es una actividad.*

El algoritmo propuesto utiliza cinco métricas para inferir redes de interacción social a partir de un registro de eventos. A continuación, se describen cada una de ellas (tomadas de (Van der Aalst y Song 2004, 2004 ; Smith y Fingar 2003; Curtis, Kellner y Over 1992; Van Der Aalst 2011)):

- **Handover of work**: Esta métrica determina quién le entrega trabajo a quién. Esta información puede ser extraída de un registro de eventos, encontrando actividades posteriores en el mismo caso, donde la primera actividad se completa con un recurso y la segunda se completa con otro recurso.
- **Subcontracting**: Esta métrica es similar a **Handover of work**. Mientras en la relación anterior entre dos individuos es unidireccional, en esta es bidireccional. Teniendo en cuenta un solo caso de un registro de

evento y dos recursos, un recurso subcontrata a otro, cuando este último ejecuta una actividad que se encuentra entre dos actividades ejecutadas por el primer recurso.

- **Working together:** Dos recursos trabajan juntos si llevan a cabo actividades en el mismo caso de un registro de eventos. Esta métrica sólo tiene en cuenta la frecuencia en que los recursos trabajan en el mismo caso.
- **Similar task:** Esta métrica determina qué recursos ejecutan el mismo tipo de actividades. Para ello, cada recurso tiene su propio perfil en función de la frecuencia en que llevan a cabo actividades específicas.
- **Reassignment:** La idea básica de esta métrica es detectar la asignación de actividades de un recurso a otro: si un recurso delega frecuentemente una actividad a otro, pero no viceversa, es probable que el primero sea jerárquicamente más alto que el segundo.

Las redes de interacción social que se obtienen con la aplicación de dichas métricas son la entrada para realizar el ARS y deben ser coherentes con las definiciones 4, 5 y 6 (tomadas de (Wasserman y Faust 1994)).

DEFINICIÓN 4 Red social: *Es un conjunto finito de actores que tienen un determinado número de relaciones definidas entre ellos. Una red social es modelada como un grafo $G = (V, E)$. En la presente investigación, el conjunto de vértices V se corresponde a los actores y el conjunto de aristas E a las relaciones sociales.*

DEFINICIÓN 5 Actor: *Es la entidad social que interactúa con otras entidades. Esta puede ser una persona, un grupo de individuos, una organización, una ciudad, una nación, entre otros.*

DEFINICIÓN 6 Vínculo o conexión: *Es el tipo de relación que conecta un par de actores distintos. Estos vínculos pueden ser de naturaleza diferente: relaciones transnacionales, de comunicación, instrumentales, sentimentales, de poder o autoridad, de parentesco, entre otros.*

A partir de las redes de interacción social inferidas se calculan tres medidas de centralidad relevantes en el ARS descritas por las definiciones 7, 8 y 9: centralidad, cercanía e intermediación (tomadas de (Beauchamp, 1965; Freeman, 1978)).

DEFINICIÓN 7 Centralidad: *Sea A la matriz de adyacencia de un grafo G , y $g(v_i)$ el grado del vértice v_i , entonces el grado de centralidad c_i^{DEG} del vértice v_i se define: $c_i^{DEG} = g(v_i)$.*

DEFINICIÓN 8 Cercanía: *Es la más conocida y utilizada de las medidas radiales de longitud. Se basa en calcular el promedio de las distancias más cortas desde un vértice hacia todos los demás. La cercanía se define como:*

$c_i^{CLO} = e_i^T S$, donde $e_i^T S$ es la transpuesta de un vector columna cuyo i -ésimo elemento es 1 y el resto son 0, S es la matriz cuyos elementos en la posición (i, j) corresponden a la distancia más corta desde el vértice v_i al vértice v_j .

DEFINICIÓN 9 Intermediación: *La intermediación es una medida que cuantifica la frecuencia o el número de veces que un vértice actúa como un puente a lo largo del camino más corto entre otros dos vértices. De manera formal, la intermediación se define como: $c_i^{BET} = \sum_{j,k} (b_{jik} / b_{jk})$, donde b_{jk} es la cantidad de caminos mínimos desde v_j a v_k y b_{jik} la cantidad de caminos mínimos desde v_j a v_k que pasan por v_i .*

Con las medidas centralidad, cercanía e intermediación calculadas se determinan los factores popularidad, eficiencia, y sobrecarga respectivamente estableciendo una relación entre ellos y siendo coherente con las definiciones 10, 11 y 12 (tomadas de (Weske, 2010)).

DEFINICIÓN 10 Popularidad: *La popularidad de un recurso en una red social se mide por el número de conexiones que tiene.*

Este concepto se puede relacionar estrechamente con la medida de centralidad. El nodo con mayor centralidad en la red es el nodo con mayor probabilidad de relacionarse con otro nodo que se adicione. Por lo que se puede inferir que el nodo con mayor centralidad, es el recurso con mayor probabilidad de participar en una actividad agregada al proceso. Además, la centralidad de un nodo en la red indica su nivel de importancia con respecto a otros nodos, determinando en una red los recursos indispensables de un proceso.

DEFINICIÓN 11 Eficiencia: *Es la forma de optimizar los flujos de comunicación y los recursos de una red de interacción social en el menor tiempo posible.*

En el ARS, la medida de cercanía permite conocer cuán eficiente es el flujo de información entre los recursos de un proceso determinado. La eficiencia dentro de un proceso puede estar condicionada por dos factores importantes: el tiempo y/o el costo de realizar una actividad del proceso (Wey et al., 2008). Estos dos factores juegan un papel fundamental a la hora de aplicar esta medida debido a que, basados en ellos es que dicha medida determina los flujos con menor tiempo y/o coste para su realización de forma ordenada. Por lo que se puede relacionar la medida de cercanía con la eficiencia, debido a que esta última se calcula haciendo uso de la primera.

DEFINICIÓN 12 Sobrecarga: *La sobrecarga de un recurso del proceso evita que dicho recurso pueda transferir la información hacia los demás recursos del proceso en el momento necesario, o sea, en un mismo instante de tiempo sus entradas van a ser mayores que sus salidas* (Barto, 2009).

Teniendo en cuenta la definición 9 se puede determinar cuál es el recurso del proceso que está haciendo función de puente entre otros. Por lo que se puede afirmar que un nodo con alta intermediación es un recurso con más sobrecarga de trabajo. A partir de este análisis, se puede identificar posibles cuellos de botella.

Finalmente se obtiene como resultado la red de interacción social asociada al registro de eventos recibido como entrada. La información derivada puede ser utilizada para comprender procesos complejos y factores que influyen en su rendimiento (Calders y Pechenizkiy 2012; Cairns et al., 2014).

Resultados y discusión

El resultado de esta investigación es un algoritmo para inferir redes de interacción social a partir de la ejecución de un proceso. En la Figura 1 se muestra el diagrama de flujo del algoritmo propuesto donde HW, SC, WT, ST y RA están asociados a las métricas Handover of work, Subcontracting, Working together, Similar task y Reassignment respectivamente, así mismo P, E y S representan los factores Popularidad, Eficiencia y Sobrecarga.

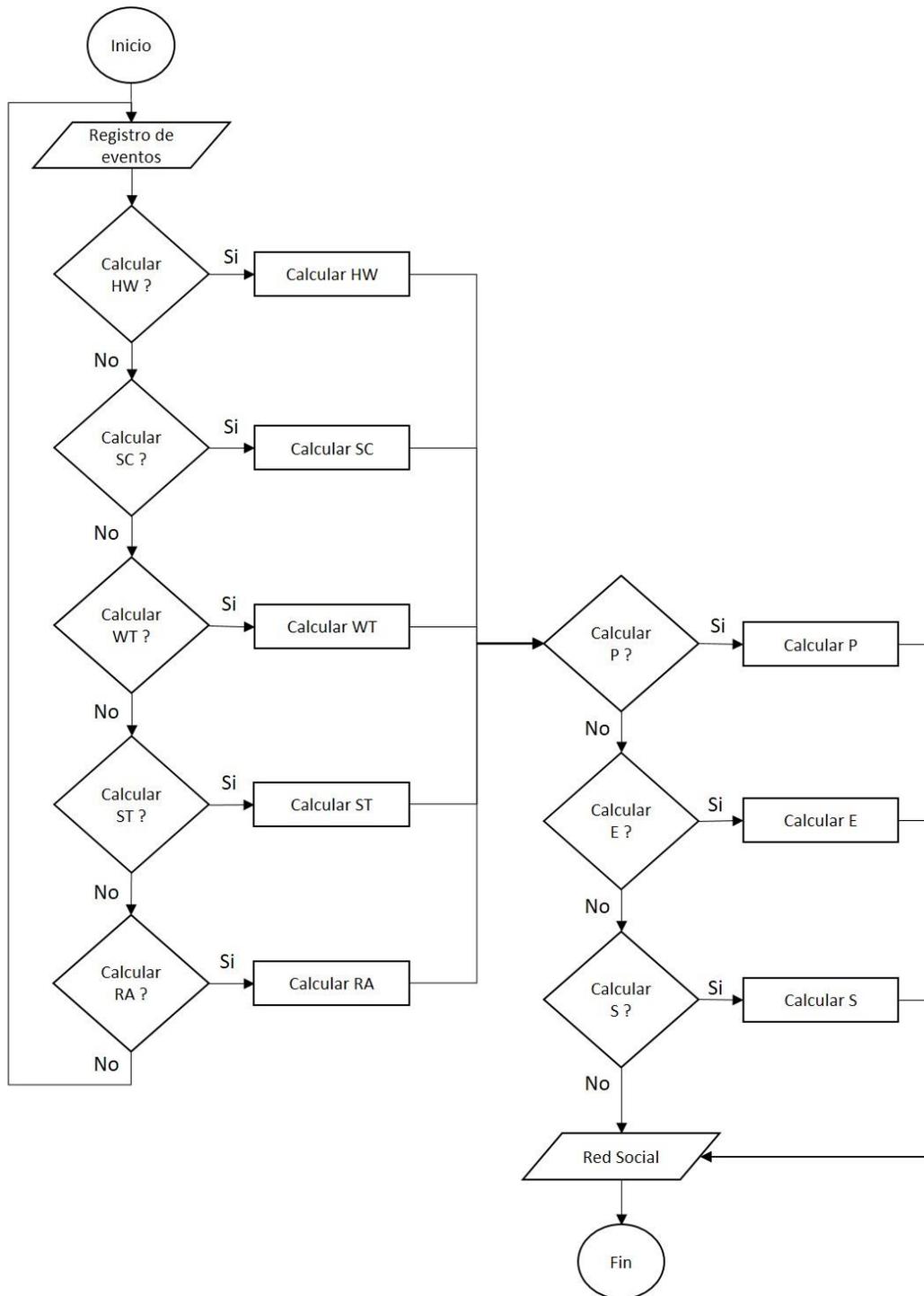


Figura 1: Diagrama de flujo del algoritmo propuesto.

Caso de estudio

Para la realización computacional del algoritmo propuesto se desarrolló un complemento para la herramienta de código abierto y con capacidad de minería de procesos ProM debido a su carencia actualmente de realizar ARS a nivel de subprocesos (Verbeek et al. 2010). Además, se utilizó un caso de estudio para comprobar la correctitud del algoritmo propuesto. En dicho caso de estudio se aplica el algoritmo para cada subproceso y se calculan los factores popularidad, eficiencia y sobrecarga con el objetivo de mejorar la comprensión del proceso global.

El caso de estudio está asociado a un registro de eventos perteneciente al registro de llamadas de uno de los servicios hoteleros que se efectúan en Cuba, el cual contiene datos reales y ha sido utilizado en (Martínez y Corella 2015; Corella, Martínez y Alfonso 2015). Los recursos que se ven involucrados en el proceso son recursos humanos que representan los recepcionistas del hotel. Las actividades que realizan cada uno de ellos representan las formas de relacionarse uno con los otros. El registro de eventos contiene 162 instancias de procesos y un total de 1042 eventos. Las redes de interacción social son inferidas por la métrica Handover of work debido a su simplicidad en los análisis.

La Figura 2 muestra la red de interacción social asociada al proceso global del caso de estudio. La visualización del proceso global permite realizar el análisis de lo general a lo particular, lo cual contribuye a disminuir la complejidad del modelo de proceso, poniéndose en evidencia uno de los factores para la comprensión del proceso. Sin embargo, el análisis sobre el proceso global se dificulta, debido a que el cálculo de los factores popularidad, eficiencia y sobrecarga se torna complejo.

Según la red inferida para el proceso global en la Figura 2, se puede apreciar que existen tres recursos con la misma popularidad (R4, R8 y R9). Como resultado del alto número de interacciones entre los recursos se ve afectado el análisis de los factores popularidad, eficiencia y sobrecarga, lo cual provoca poca comprensión del proceso global.

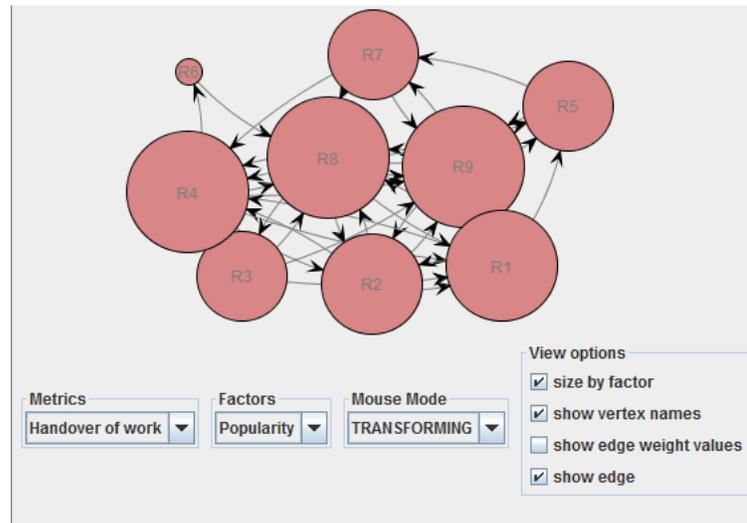


Figura 2. Red de interacción social asociada al proceso global.

Para lograr una mayor comprensión del proceso global se utilizó el método de descomposición de procesos propuesto por (Pérez et al. 2015), donde cada subproceso está asociado a un patrón de descomposición (Secuencia, Paralelismo, Selección exclusiva y Lazo). La Figura 3 muestra la red de interacción social inferida para el subproceso asociado a la primera ocurrencia del patrón Lazo. Se puede apreciar que los recursos más populares del subproceso son R4, R8 y R9, siendo recursos indispensables para el subproceso y los que mayor probabilidad tienen de extender el modelo de proceso en el futuro.

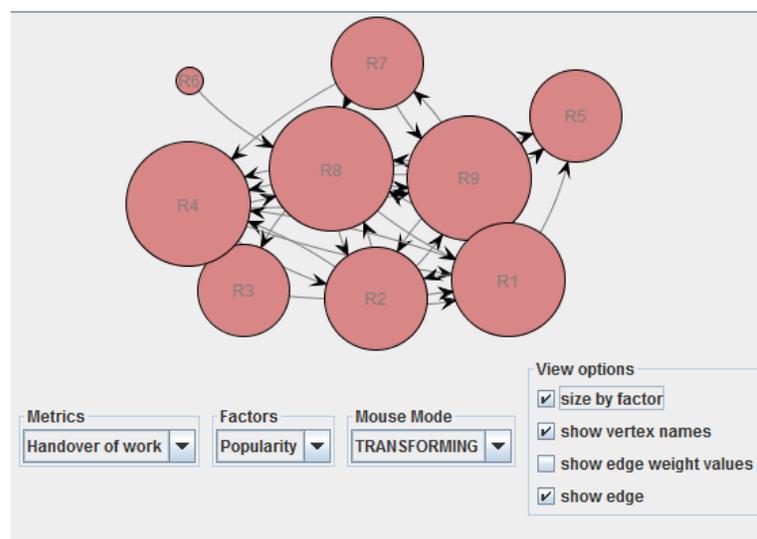


Figura 3. Cálculo de popularidad para el subproceso del primer Lazo.

La Figura 4 muestra que los recursos R1, R2, R3, R4, R6 y R8 son los de mayor eficiencia en el subproceso. Esta información brinda la posibilidad de optimizar los recursos del subproceso y conocer sus flujos de trabajo óptimos para reducir el tiempo de ejecución del subproceso.

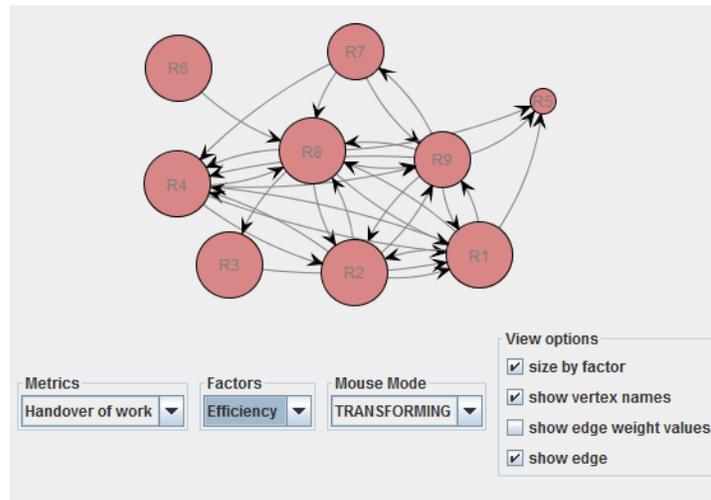


Figura 4. Cálculo de eficiencia para el subproceso del primer Lazo.

La Figura 5 muestra el cálculo del factor sobrecarga para el subproceso. Como se puede distinguir en la red, el recurso con mayor sobrecarga de trabajo es R8. Este dato es de vital importancia en el ARS porque puede que en algún momento este recurso sea el causante de un cuello de botella en el flujo de trabajo del subproceso.

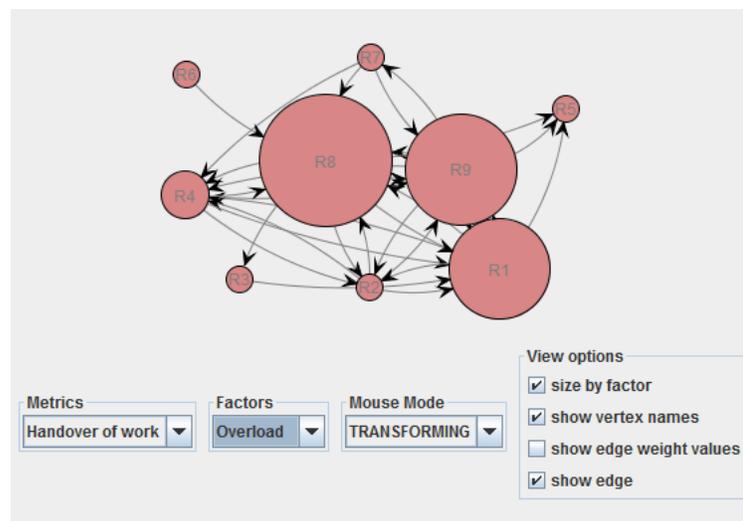


Figura 5. Cálculo de sobrecarga para el subproceso del primer Lazo.

El algoritmo se aplicó para cada subproceso, obteniendo una especificación más detallada del proceso global. El penúltimo subproceso está asociado a la primera ocurrencia del patrón Selección exclusiva. En esta iteración se puede apreciar que resulta más fácil realizar el ARS con el desglose de cada subproceso que componen al proceso global. En la Figura 6 se muestran los nodos con mayor popularidad (R1, R8 y R9), lo cual implica que son recursos indispensables y elementos a tener en cuenta si se decide engrandecer el proceso.

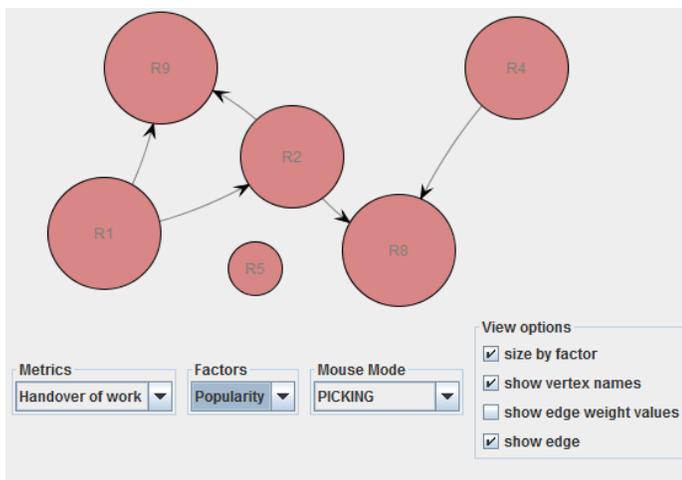


Figura 6. Cálculo de popularidad para el subproceso de la primera Selección exclusiva.

En la red de interacción social de la Figura 7 se muestran los recursos más eficientes del subproceso asociado (R1, R2 y R4), siendo elementos a tener en cuenta en la optimización del proceso.

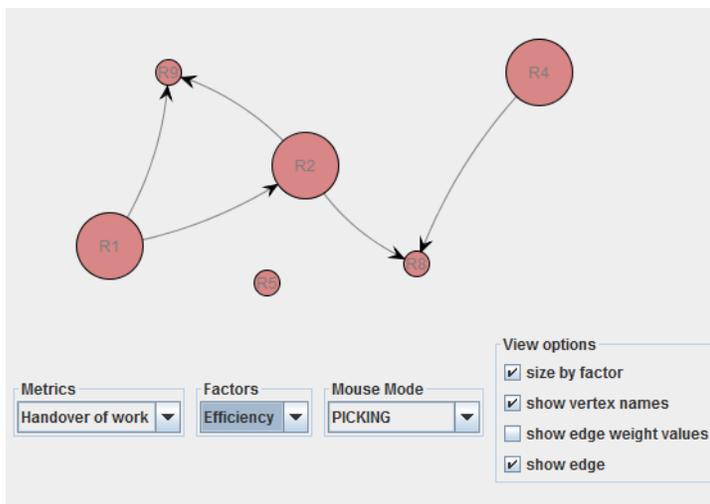


Figura 7. Cálculo de eficiencia para el subproceso de la primera Selección exclusiva.

En la Figura 8 se muestra la red de interacción social del subproceso asociado. Se puede apreciar que el recurso R2 es aquel con mayor sobrecarga, lo que significa que es el recurso que más actúa como intermediario entre otros dos y pudiera ser el causante de un cuello de botella.

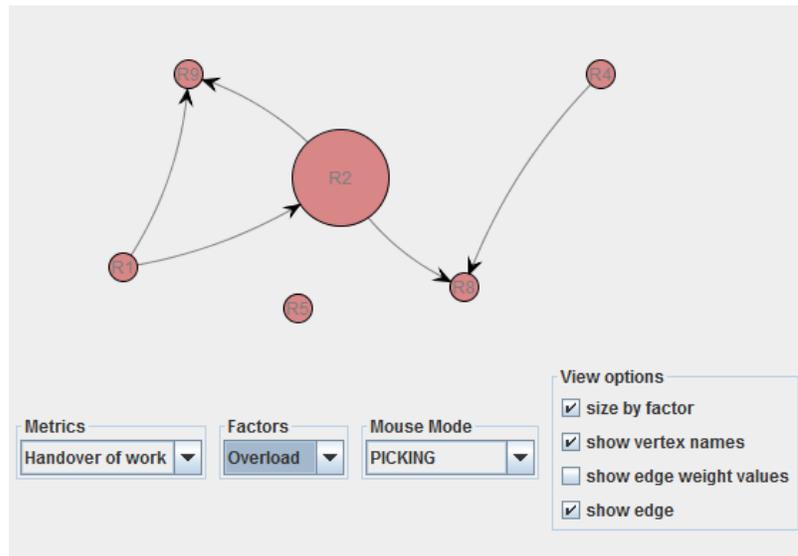


Figura 8. Cálculo de sobrecarga para el subproceso de la primera Selección exclusiva.

Por todo lo anterior se puede afirmar que el resultado obtenido al aplicar el algoritmo propuesto permitió el cálculo de los factores popularidad, eficiencia y sobrecarga, a partir de las redes de interacción social inferidas para los subprocesos que componen a un proceso complejo. La propuesta presentada garantiza una mejora en la comprensión de procesos complejos debido a que los análisis pueden realizarse a nivel de subprocesos, no siendo así con los complementos que presenta ProM para SNM que sólo se puede analizar el proceso global.

Conclusiones

Como resultado de la presente investigación se obtuvo un algoritmo para la inferencia de redes de interacción social a partir de eventos con alto número de interacciones entre los recursos, mejorando la comprensión de procesos complejos en su dimensión social; el mismo garantiza extensibilidad. La extensibilidad del algoritmo presentado radica en la posibilidad de adicionar más factores para el ARS sin afectar los demás componentes del mismo, enriqueciendo de esta forma los análisis posteriores.

Las métricas existentes para inferir redes de interacción social a partir de subprocesos, permiten obtener información relevante acerca de las interacciones de los recursos en un proceso determinado. Así mismo, el cálculo de las medidas

de centralidad en redes de interacción social permite realizar análisis del comportamiento de los recursos en un proceso determinado; sin embargo, la comprensión de dicho proceso se ve dificultada cuando existe un alto número de interacciones entre sus recursos, provocando que el análisis de dichas medidas también se vea afectado.

La realización del caso de estudio evidenció la efectividad del algoritmo propuesto, mostrando que dicho algoritmo permite identificar recursos indispensables en un proceso, flujos de trabajo óptimos, recursos con mayor probabilidad para extender el modelo de proceso y recursos con sobrecarga de trabajo. A partir de los resultados obtenidos se justifica el uso del algoritmo propuesto en cualquier entorno donde se tengan las trazas de la ejecución de un proceso, con el objetivo de inferir redes de interacción social a partir de eventos con alto número de interacciones entre los recursos, facilitando la comprensión de procesos complejos en su dimensión social.

Referencias

- AALST, W., 2011. *Process mining: Discovery, conformance and enhancement of business processes*. S.l.: Springer-Verlag, Berlin.
- AGRAWAL, R., GUNOPULOS, D. y LEYMAN, F., 1998. Mining process models from workflow logs. [en línea]. S.l.: Springer, pp. 467–483. Disponible en: <http://link.springer.com/chapter/10.1007/BFb0101003>.
- BARTO, A.G., 2009. Skill characterization based on betweenness. [en línea]. S.l.: s.n., pp. 1497–1504. Disponible en: <https://papers.nips.cc/paper/3411-skill-characterization-based-on-betweenness.pdf>.
- BEAUCHAMP, M.A., 1965. An improved index of centrality., vol. 10, no. 2, pp. 161–163.
- CAIRNS, A.H., GUENI, B., FHIMA, M., CAIRNS, A., DAVID, S. y KHELIFA, N., 2014. Towards Custom-Designed Professional Training Contents and Curriculums through Educational Process Mining. , pp. 53–58.
- CALDERS, T. y PECHENIZKIY, M., 2012. Introduction to the special section on educational data mining. , vol. 13, no. 2, pp. 3–6.
- CORELLA, A.B., MARTÍNEZ, Y.A. y ALFONSO, D.P., 2015. Componente para la extracción de registros de eventos en formato XES del sistema ZUN Suite. *Serie Científica de la Universidad de las Ciencias Informáticas*, vol. 8, no. 2.
- CURTIS, B., KELLNER, M.I. y OVER, J., 1992. Process modeling., vol. 35, no. 9, pp. 75–90.
- FREEMAN, L.C., 1978. Centrality in social networks conceptual clarification., vol. 1, no. 3, pp. 215–239.

- MARTÍNEZ, Y.A. y CORELLA, A.B., 2015. Analysis of event logs extracted in XES format from hotel management system ZUN Suite.,
- PÉREZ, D., 2015. *Método para el diagnóstico de procesos de negocio a partir de registros de eventos con ruido y ausencia de información*. La Habana: Universidad de las Ciencias Informáticas.
- PÉREZ, D., YZQUIERDO HERRERA, R., PUPO HERNÁNDEZ, E. y LÓPEZ JIMÉNEZ, R., 2015. Algoritmo para la identificación de variantes de procesos., vol. 9, no. 4, pp. 199–215.
- PÉREZ, D., 2006. *Contribución de las tecnologías de la información a la generación de valor en las organizaciones: un modelo de análisis y valoración desde la gestión del conocimiento, la productividad y la excelencia en la gestión* [en línea]. S.l.: s.n. Disponible en: <http://repositorio.unican.es/xmlui/handle/10902/1173>.
- SMITH, H. y FINGAR, P., 2003. *Business process management: the third wave* [en línea]. S.l.: s.n. Disponible en: <https://pdfs.semanticscholar.org/ea5a/1375811799a5be2bb2c696023d4c43bc6e5d.pdf>.
- VAN DER AALST, W., 2011. Manifiesto sobre Minería de Procesos.,
- VAN DER AALST, W., ADRIANSYAH, A. y VAN DONGEN, B., 2012. Replaying history on process models for conformance checking and performance analysis., vol. 2, no. 2, pp. 182–192.
- VAN DER AALST, W., WEIJTERS, T. y MARUSTER, L., 2004. Workflow mining: Discovering process models from event logs. , vol. 16, no. 9, pp. 1128–1142.
- VAN DER AALST, W.M. y SONG, M., 2004. Mining social networks: Uncovering interaction patterns in business processes. [en línea]. S.l.: s.n., pp. 244–260. Disponible en: http://link.springer.com/10.1007/978-3-540-25970-1_16.
- VERBEEK, H., BUIJS, J., VAN DONGEN, B.F. y VAN DER AALST, W., 2010. Xes, xesame, and prom 6. [en línea]. S.l.: s.n., pp. 60–75. Disponible en: http://link.springer.com/chapter/10.1007/978-3-642-17722-4_5.
- WASSERMAN, S. y FAUST, K., 1994. *Social network analysis: Methods and applications* [en línea]. S.l.: s.n. Disponible en: <https://books.google.com/books?hl=es&lr=&id=CAm2DpIqRUIC&oi=fnd&pg=PR21&dq=Social+network+analysis:+Methods+and+applications.+S.l.:+Cambridge+university+press&ots=HvOswe0zR9&sig=BZG121eBqXWGkTgewIrAGQmUn-Y>.

WESKE, M., 2010. *Business process management: concepts, languages, architectures* [en línea]. S.l.: s.n. Disponible en: <http://dl.acm.org/citation.cfm?id=1951579>.

WEY, T., BLUMSTEIN, D.T., SHEN, W. y JORDÁN, F., 2008. Social network analysis of animal behaviour: a promising tool for the study of sociality., vol. 75, no. 2, pp. 333–344.