

Diseño de Redes de Sensores Utilizando Modelos Gráficos Probabilísticos

Azulilen Foricher¹, José Hernández¹, Mercedes Carnero¹, Mabel Sánchez²

¹ Grupo de Optimización, Facultad de Ingeniería, Universidad Nacional de Río Cuarto, Río Cuarto, Argentina

{aforicher, jlh, mcarnero}@ing.unrc.edu.ar

² Planta Piloto de Ingeniería Química, UNS- CONICET, Bahía Blanca, Argentina
msanchez@plapiqui.edu.ar

Abstract. En este trabajo se aborda el diseño óptimo de redes de sensores para plantas químicas utilizando estrategias de optimización estocásticas. El problema consiste en seleccionar el tipo, número y ubicación de los nuevos sensores que proporcionen la cantidad y calidad necesaria de la información requerida del proceso. En el diseño de redes de sensores la decisión importante que debe hacerse con respecto a cada variable de flujo es si ésta se mide o no. Para formular matemáticamente estas decisiones, se emplean variables binarias las cuales indican la presencia o ausencia de sensores. El problema suele ser multimodal e involucra, en casos reales, un gran número de variables binarias, por lo que debe ser resuelto un problema de optimización combinatoria sujeto a restricciones de gran tamaño. En estos casos, es muy valioso contar con un procedimiento de solución que proporcione, al menos, una buena solución, sino el óptimo global, y que además se pueda ejecutar en ordenadores paralelos para reducir los tiempos de ejecución. En este trabajo se aplican nuevas estrategias para resolver el problema de localización óptima de sensores basadas en los algoritmos de estimación de distribuciones que hacen uso de un modelo gráfico probabilístico, aprendido a partir del conjunto de soluciones más prometedoras. Se proporcionan las características distintivas de las metodologías propuestas así como su desempeño en la resolución de diferentes diseños de redes de instrumentación extraídos de la literatura.

Keywords: Redes de sensores; algoritmos evolutivos; algoritmos de estimación de distribuciones.

1 Introduction

En la industria de procesos es imprescindible contar con información exacta y precisa para llevar adelante actividades tales como: optimización en línea, planeamiento, control estadístico de procesos, diagnóstico de fallas, mantenimiento predictivo, etc. Dado que la información proviene de los instrumentos instalados en la planta, el diseño óptimo de la red de sensores resulta esencial para satisfacer los objetivos de la empresa.

El planeamiento de un Sistema de Instrumentación comprende la definición de los objetivos globales, la selección de las variables medidas y la especificación de los detalles de implementación, tales como el intervalo entre mediciones, el método de toma de muestra, la interfaz con el operador, etc.

El diseño de la red de sensores de una planta de procesos con fines de monitoreo consiste en determinar para cada variable de proceso si será medida o no. Habitualmente esta decisión se basa en experiencias previas con plantas similares o reglas empíricas. Sin embargo es posible modelar este problema como un problema de optimización discreta del tipo NP-completo. La resolución del mismo permite concretar el diseño y/o actualización del sistema de instrumentación de manera eficiente y automática.

En procesos que involucran un gran número de variables, aparecen las heurísticas como una herramienta de optimización importante. Entre éstas, los algoritmos evolutivos, AE, en sus diferentes variantes, han sido utilizados extensamente durante los últimos años para resolver precisamente problemas de optimización de tipo combinatorio. Dentro de este paradigma se han comenzado a aplicar los algoritmos de estimación de distribuciones, AED, más recientemente. Si bien estos comparten con los AEs su carácter de algoritmos poblacionales que evolucionan iterativamente, en los AEDs la búsqueda está centrada en encontrar una distribución de probabilidades conjunta que al ser muestreada genere una población de soluciones de alta calidad, [1], [2]. En el presente trabajo se propone resolver el diseño general de una red de sensores utilizando dos tipos diferentes de AEDs con el fin de comparar el desempeño de ambos entre sí.

2 Diseño General de Redes de Sensores

El diseño óptimo de una red de sensores de tipo general consiste en determinar la localización de un conjunto de instrumentos en el proceso, que satisfaga un dado criterio de optimización y permita disponer de información de la planta con los niveles de calidad y cantidad requeridos.

En particular, el diseño de una red de sensores de mínimo costo de adquisición que satisfice las restricciones de estimabilidad y precisión sobre un determinado conjunto de variables claves se formula como sigue:

$$\begin{aligned}
 & \text{Min} && \mathbf{c}^T \mathbf{q} && (1) \\
 & \text{s.t.} && && \\
 & && \hat{\sigma}_j(\mathbf{q}) \leq \sigma_j^* && \forall j \in S_\sigma \\
 & && E_k(\mathbf{q}) \geq 1 && \forall k \in S_E \\
 & && \mathbf{q} \in \{0,1\}^n &&
 \end{aligned}$$

donde \mathbf{c}^T es el vector de costos de adquisición, $\hat{\sigma}_j$ indica el desvío estándar de la estimación de la variable j obtenida después de aplicar un procedimiento de reconciliación de datos, E_k representa el grado de estimabilidad de la variable k , y q_i es una variable binaria que asume el valor 1 si la i -ésima variable es medida. Las

restricciones de precisión y estimabilidad se imponen sobre los flujos pertenecientes a los subconjuntos S_σ y S_E respectivamente, [3]. Se supone que el proceso opera en estado estacionario y que la interrelación entre las variables del proceso viene descrita por un modelo lineal. Asimismo se considera que las mediciones de las variables están afectadas por errores aleatorios, no correlacionados, normalmente distribuidos con media cero y varianza coincidente con la varianza inherente al instrumento de medición.

Se propone, para el tipo de diseño planteado en (1), analizar el comportamiento de las heurísticas que se detallan a continuación, en términos de calidad de las soluciones obtenidas y tiempos de cómputo insumidos con miras a su implementación para la resolución de problemas de diseño óptimo de redes de sensores en grandes plantas químicas.

3 Algoritmos Propuestos

Los AEDs pueden ser clasificados dentro del conjunto de técnicas denominadas Algoritmos Evolutivos. Al igual que estos, son procedimientos estocásticos que generan una población de individuos $D^t = \{\mathbf{q}_1^t, \dots, \mathbf{q}_M^t\}$ para cada iteración t , los cuales son posibles soluciones del problema y se representan mediante una estructura de datos \mathbf{q} . Cada individuo \mathbf{q}_i^t se evalúa utilizando una medida de su aptitud o función de *fitness*. Se distinguen de los AEs principalmente en la forma de guiar la búsqueda de soluciones óptimas. Mientras que estos últimos utilizan información de los individuos para realizar las operaciones de mutación y cruzamiento, los AEDs construyen y evolucionan un modelo probabilístico que mejor represente el espacio de soluciones. Los datos obtenidos luego de la evaluación de cada individuo se utilizan para actualizar el modelo probabilístico y a partir de éste se muestrea la siguiente población. En este contexto entonces, los individuos se consideran como potenciales soluciones al problema de optimización, y también como realizaciones de un vector aleatorio cuya función de distribución de probabilidades conjunta se ha estimado.

El algoritmo se ejecuta un cierto número de generaciones ó hasta que se satisface algún criterio de detención.

Cabe destacar que la capacidad de los AEDs de construir modelos que pueden captar las interrelaciones entre las variables y de este modo guiar la búsqueda hacia las regiones más promisorias constituye la principal fortaleza de estos métodos para resolver problemas de optimización complejos, [4]. Por otra parte el aprendizaje del modelo de distribución de probabilidades subyacente, esto es la distribución conjunta de probabilidades, es a su vez la componente más importante y más compleja de los AED, [5].

Dependiendo de la forma de construir el modelo probabilístico, los numerosos AEDs que se encuentran en la literatura se distinguen entre modelos univariados y multivariados. Los primeros son simples porque asumen que las variables del problema son independientes. En cambio los segundos son complejos dado que buscan percibir y modelar las interrelaciones presentes entre las variables.

En este trabajo se analizan algoritmos de resolución multivariados. Se han considerado dos exponentes de esta categoría que difieren en las hipótesis principales acerca de la estructura de dependencias entre variables. A continuación se describen las características principales de cada uno de ellos.

3.1 Algoritmo BN, Bayesian Networks

Los modelos gráficos probabilísticos, [6], [7], son herramientas utilizadas para representar interacciones en sistemas complejos en términos de dependencias probabilísticas. De manera general constan de dos elementos: una estructura que refleja las dependencias entre variables y una colección de valores numéricos que indican el grado de las interacciones. Estos parámetros son usualmente probabilidades condicionales. Un modelo gráfico probabilístico se representa mediante un grafo G , en el que cada nodo está asociado a una variable aleatoria y los enlaces entre nodos muestran las dependencias condicionales entre las variables. Si la estructura de la red es un grafo acíclico dirigido, éste simboliza una factorización de la distribución conjunta $p(\mathbf{q})$, para el vector aleatorio $\mathbf{Q}=(Q_1, \dots, Q_n)$. Este tipo de modelo se conoce como red bayesiana. La implementación de redes bayesianas en AED involucra el aprendizaje de la estructura, la determinación del conjunto de parámetros y la posterior simulación del modelo aprendido. Los diferentes enfoques encontrados en la literatura acerca de esta metodología en AEDs difieren esencialmente en la forma de realizar algunos de estos pasos.

En este trabajo se implementó un algoritmo que utiliza redes bayesianas, algoritmo BN, con las siguientes características: población D_0 inicializada mediante muestreo utilizando una distribución uniforme, selección por truncamiento con umbral igual a 0.5, reemplazo de los 10 mejores individuos de la generación anterior en la población corriente y muestreo del modelo mediante *Probabilistic Logic Sampling*, PLS, [8]. El modelo probabilístico se aprende mediante el algoritmo B, [9] y el Criterio de Información Bayesiano, BIC, [10], como métrica para determinar el grado de ajuste.

3.2 Algoritmo LMP, Learning Marginal Products

La descomposición de la distribución conjunta de probabilidades de un vector aleatorio en un producto de múltiples distribuciones de probabilidades se denomina factorización de productos marginales. El conjunto de variables se particiona en subconjuntos disjuntos de diferentes tamaños, es decir que cada variable aparece en sólo un subconjunto. Para cada uno de estos grupos se considera una distribución multivariada, cuya tabla de frecuencias debe ser computada. Esta factorización de la distribución conjunta de probabilidades no modela relaciones de dependencia entre variables que pertenecen a diferentes grupos.

En contraste con Redes Bayesianas, no se utilizan distribuciones condicionales en la factorización de la distribución conjunta sino que se supone que todas las variables que aparecen en un factor se distribuyen de manera conjunta. El número de parámetros a estimar, esto es el tamaño de las tablas de frecuencia asociadas a cada factor, crece de manera exponencial con la cardinalidad de cada subconjunto de la

partición propuesta. Este concepto de factorización de la distribución conjunta se utiliza en el algoritmo LMP para realizar el aprendizaje estructural del modelo probabilístico. En este trabajo se ha implementado el algoritmo LMP desarrollado por Santana et al. [11] que utiliza el algoritmo de propagación de afinidades (AffEDA) para llevar a cabo la partición del conjunto de variables en subconjuntos mutuamente excluyentes. Este es un algoritmo de agrupamiento, (*clustering*): dado un conjunto de puntos y una medida de similitud entre ellos, proporciona grupos de puntos similares y además para cada grupo da un ejemplar representativo. La medida de similitud es la información mutua entre cada par de variables aleatorias, la cual expresa la información que ellas comparten.

El número de grupos o *clusters* no se determina de antemano, sino que es entregado por el algoritmo de propagación de afinidades. El tamaño de cada cluster determina obviamente el número de factores que aparecen en cada distribución marginal y por lo tanto el número de parámetros a estimar.

4 Experimentos y Resultados

Para analizar el desempeño de las estrategias de resolución planteadas se tomó un ejemplo conocido en la literatura [12] consistente en un sistema de vapor de una planta de producción de metanol que cuenta con 28 corrientes y 11 equipos. La Fig. 1 muestra el diagrama de flujo del sistema mencionado.

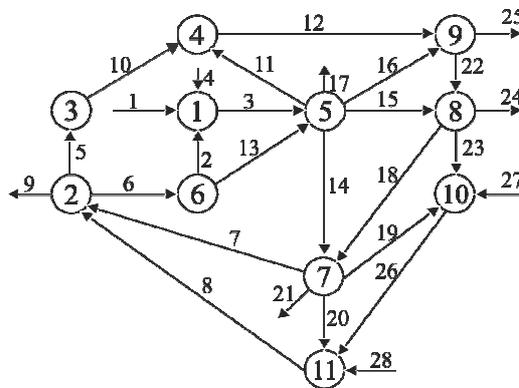


Fig. 1. Diagrama de flujo del sistema de vapor de una planta de metanol

Se plantearon dos instancias del mismo problema, llamados diseño A y diseño B. La tabla 1 muestra las restricciones de estimabilidad y precisión impuestas a cada diseño. Se supone que las variables están relacionadas exclusivamente a través de las ecuaciones de balance de masa y que el desvío estándar de los caudalímetros disponibles corresponde al 2% del valor verdadero del caudal medido.

Tabla 1. Restricciones del problema de diseño

Caso	Restricciones
A	$E \geq 1$ para las corrientes 4, 8, 17, 21, 23, 25 $\sigma_4^* = 2.20, \sigma_8^* = 3.28, \sigma_{21}^* = 1.75, \sigma_{25}^* = 1.71$
B	$E \geq 1$ para las corrientes 4, 5, 7, 8, 12, 16, 18, 20, 27,28 $\sigma_4^* = 2.20, \sigma_5^* = 1.07, \sigma_8^* = 3.28, \sigma_{12}^* = 1.35, \sigma_{27}^* = 1.42, \sigma_{28}^* = 1.45$

En este trabajo se analiza la capacidad de los algoritmos para encontrar buenas soluciones al problema de optimización planteado. Dicha capacidad está relacionada con la calidad de la estimación del modelo probabilístico subyacente, la cual depende de manera directa del tamaño de la población del algoritmo. Por tal motivo, el análisis se basa en los resultados obtenidos al resolver los dos problemas de diseño previamente citados con ambos algoritmos, utilizando tres tamaños diferentes de población. Para cada uno de estos experimentos se realizaron cien ejecuciones independientes sobre una computadora equipada con un Procesador Intel® Core (TM) i7 CPU 920 @ 2.67 GHz, 6GB RAM, usando MatLab Release 14. El criterio de detención utilizado fue fijar un número de iteraciones igual a 200 en cada una de las corridas.

La capacidad de los algoritmos para encontrar buenas soluciones se evalúa en función de sus diferencias con los óptimos de referencia, tomados de la literatura, [13]. Para el diseño A este valor es de 752,26, el cual corresponde a situar sensores en las corrientes 1 4 6 7 9 10 11 14 16-24. En el caso del diseño B el óptimo de referencia corresponde a un valor de 1178,06 e implica medir las variables 1, 2, 4, 5-7, 9-11, 13, 15-24, 26-28.

En la figura 2 puede apreciarse el comportamiento de las estrategias BN y LMP en la resolución del diseño A.

Se grafican los mejores valores encontrados en cada una de las 100 corridas ordenados de menor a mayor, es decir desde mejores a peores soluciones, para cada algoritmo y para tamaños de población inicial de 50, 100 y 200 individuos.

Se observa claramente que el desempeño del algoritmo LMP es superior al del BN, ya que la totalidad de los mejores valores encontrados por él se halla concentrada alrededor del mínimo de referencia, mientras que los valores obtenidos por el BN están apartados de éste en la mayoría de los casos. Se distingue también que la influencia del número de individuos que forman la población inicial es mayor para el algoritmo BN que para el LMP, debido a que el conjunto de puntos representativos de los mejores valores para cada tamaño están distribuidos de manera dispersa.

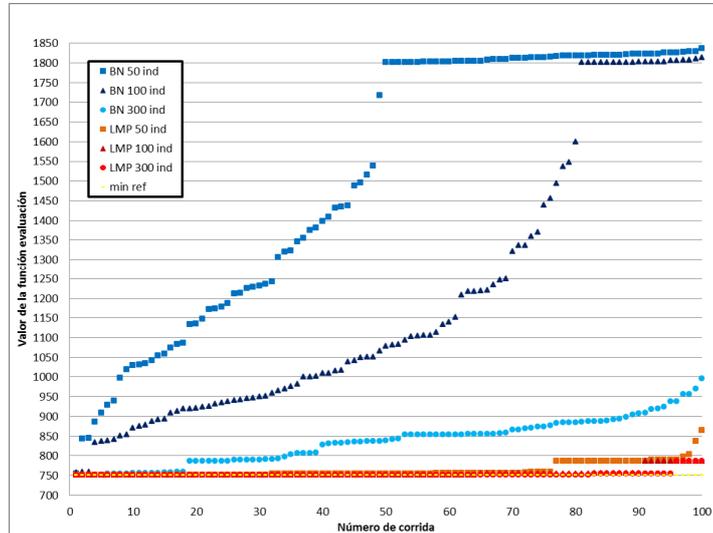


Fig. 2. Evolución de los algoritmos BN y LMP para el diseño A

Dado que la variación del tamaño de la población tiene un impacto directo sobre los tiempos de cómputo, se analizaron también los diferentes valores de tiempo de proceso para cada caso. En la figura 3 se observan los tiempos totales de ejecución de los dos algoritmos mencionados al variar el tamaño de la población inicial. Cada punto en el gráfico representa el tiempo insumido para llevar a cabo 100 ejecuciones con 200 iteraciones cada una para un tamaño particular de población.

El algoritmo LMP requirió menor tiempo para encontrar los mejores valores en todos los casos. Con esta última observación se puede aseverar que, para el diseño A, es más conveniente el algoritmo LMP, tanto respecto a la calidad de las soluciones como al tiempo empleado en encontrarlas.

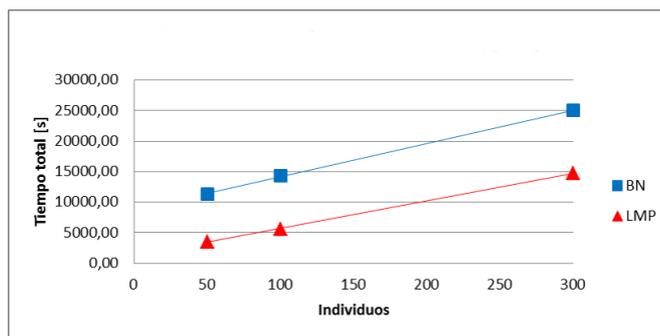


Fig. 3. Comparación de tiempo de ejecución para el diseño A

Se repitieron los experimentos para el diseño B y se plasmaron los mejores valores encontrados en las 100 ejecuciones con cada algoritmo y para los 3 tamaños de

población. La figura 4 muestra las soluciones obtenidas, y en la figura 5 se grafican los tiempos totales de ejecución.

En la figura 4 los resultados se ordenaron siguiendo un orden decreciente en cuanto a la calidad de la solución. El conjunto de puntos obtenidos por el algoritmo LMP se visualiza siempre por debajo del obtenido por el algoritmo BN, y más próximo al mínimo de referencia, por lo cual se considera que el primero se comporta mejor.

También se puede concluir que el algoritmo BN posee un pobre desempeño ya que la gran mayoría de los mejores valores encontrados por él se alejan del mínimo de referencia. Esto también se aprecia para los puntos del algoritmo LMP con 50 individuos en su población inicial.

En relación a los diferentes tamaños de población, se comprueba que los algoritmos convergen a una solución de mejor calidad cuanto mayor es el número de individuos utilizado, corroborando la predicción realizada en párrafos anteriores. Un modelo probabilístico que tiene una muestra de mayor tamaño es capaz de producir mejores estimaciones. En el algoritmo BN esta influencia es más notoria ya que las redes Bayesianas son modelos gráficos probabilísticos de mayor complejidad, que demandan muestras de mayor tamaño que los utilizados por LMP.

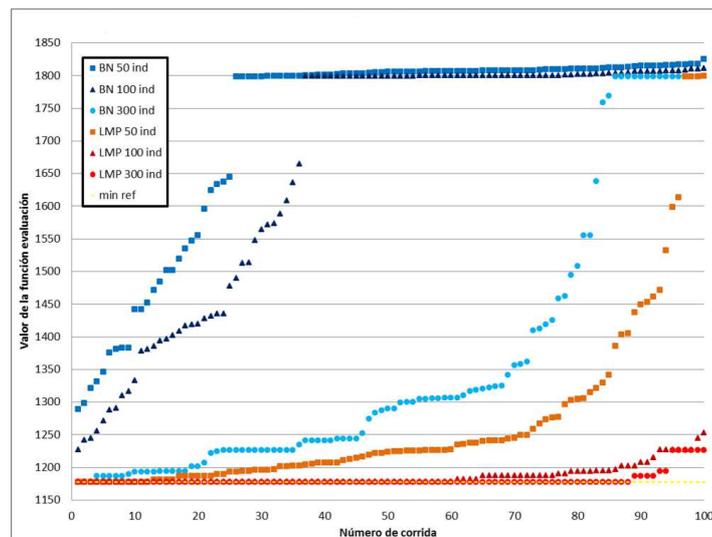


Fig. 4. Evolución de los algoritmos BN y LMP para el diseño B

La figura 5 muestra que, al igual que en el caso anterior, la variación del tiempo de ejecución es lineal respecto del aumento de la cantidad de individuos utilizada por el método de optimización y que el algoritmo LMP tiene un proceder más rápido que el BN al visualizarse la recta que lo representa por debajo de la correspondiente al BN. Finalmente se puede concluir que, para el diseño B, es conveniente aplicar el algoritmo LMP, dado que su desempeño es superior al de BN tanto en la calidad de las soluciones como en el tiempo de ejecución.

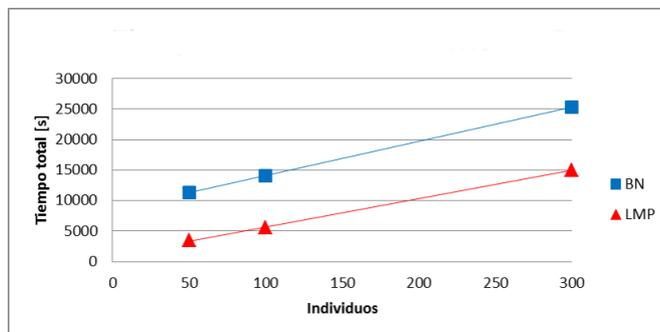


Fig. 5. Comparación de tiempo de ejecución para el diseño B

5 Conclusiones

En este trabajo se ha presentado la aplicación de metaheurísticas basadas en algoritmos de estimación de distribuciones de orden superior para la resolución del diseño óptimo de redes de instrumentación en plantas de proceso. Se implementaron dos propuestas diferentes para la fase de aprendizaje estructural de los AEDs. En un caso se utilizó el concepto de redes bayesianas para generar la función de distribución conjunta, mientras que en el otro se estimó esta función como un producto de distribuciones marginales.

Se comprobó que el desempeño de ambos algoritmos mejora sustancialmente con el aumento del tamaño poblacional. Esto se debe a que los dos procedimientos deben realizar tanto un aprendizaje estructural como paramétrico y la calidad de la estimación mejora con el incremento del número de individuos de la población. Sin embargo esto tiene un impacto directo sobre el tiempo de cómputo, mucho más acentuado en el algoritmo BN, lo cual lo hace poco apto para problemas mayores que el descrito en este trabajo.

Se probaron sobre un mismo problema dos instancias, diseños, diferentes. En ambos casos, el desempeño del algoritmo LMP es claramente superior al que muestra el algoritmo BN para todos los tamaños de población considerados. Esto se aprecia de manera evidente en todos los valores obtenidos para los estadísticos calculados para medir el desempeño.

Como conclusión general puede decirse que el algoritmo LMP, con un tamaño poblacional adecuado, puede ser utilizado como una herramienta de optimización competitiva ya que los mejores valores obtenidos corresponden a los mínimos de referencia conocidos. Trabajos futuros contemplan el análisis de la escalabilidad del algoritmo LMP además de la hibridización de éste con una estrategia de búsqueda local.

Referencias

1. P. Larrañaga, and J.A. Lozano, A review on EDAs, in Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation. Norwell, MA: Kluwer, 2002, ch 3.
2. M. Carnero, J. Hernández, M. Sánchez. A new metaheuristic based approach for the design of sensor networks, Computers and Chemical Engineering, 2013; vol. 55, pp. 83-96.
3. M. Bagajewicz, M. Sánchez, "Cost optimal design and upgrade of non-redundant and redundant linear sensor networks," AIChE J. 1999; 45: 1927-38.
4. J. Lozano, P. Larrañaga, I. Inza, E. Bengoetxea. Edas Towards a New Evolutionary Computation: Advances on Estimation of Distribution Algorithms. Springer-Verlag; 2006.
5. M. Hauschild, and M. Pelikan, "An introduction and survey of estimation of distribution algorithms," Swarm and Evolutionary Computation, vol. 1, pp. 111-128, 2011.
6. R. Howard, y J. Matheson. Influence diagrams. R. Howard, y J. Matheson editors, Readings on the Principles and Applications of Decision Analysis, vol. 2, pp. 721-764. Strategic Decision Group, Menlo Park, California, 1981.
7. S. L. Lauritzen. Graphical Models. Oxford University Press, 1996.
8. M. Henrion, "Propagating uncertainty in Bayesian networks by probabilistic logic sampling. Uncertainty in Artificial Intelligence," J. F. Lemmer and L. N. Kanal eds., North-Holland, Amsterdam, 1998, vol. 2, pp. 149-163.
9. W. Buntine, "Theory refinement on Bayesian networks," in Proc. 7th Conf. Uncertainty Artif. Intell., pp. 52-60, 1991.
10. G. Schwarz, "Estimating the dimension of a model. Annals of Statistics, 7(2), pp. 461-464, 1978.
11. R. Santana, P. Larrañaga, J. Lozano, "Learning Factorizations in Estimation of Distribution Algorithms Using Affinity Propagation." Evolutionary Computation 18(4): pp. 515-546, 2010.
12. R. Serth, and W. Heenan, "Gross Error Detection and Data Reconciliation in Steam-Metering Systems," AIChE J., vol. 32, pp. 733-741 (1986).
13. M. Carnero, J. Hernandez, M. Sanchez, "Design of Sensor Networks for Chemical Plants based on Metaheuristics ALGORITHMS ," vol. 2, N° 1, pp. 259-281, 2009