

# Metaheurísticas aplicadas al monitoreo en plantas químicas

Carlos Bermudez<sup>1</sup>, Gabriela Minetti<sup>1</sup>, Carolina Salto<sup>1</sup>

José Luis Hernandez<sup>2</sup>, Mercedes Carnero<sup>2</sup>,

Laboratorio de Investigación en Sistemas Inteligentes (LISI)

<sup>1</sup>Facultad de Ingeniería - Universidad Nacional de La Pampa

e-mail: <sup>1</sup>{bermudezc, minettig, saltoc, @ing.unlpam.edu.ar}

<sup>2</sup>Facultad de Ingeniería - Universidad Nacional de Río Cuarto

e-mail: <sup>2</sup>{jlh, mcarnero@ing.unrc.edu.ar}

**Resumen** Esta línea de investigación se enfoca en la resolución del problema de diseño de la red de sensores en las plantas de procesos, incluyendo la determinación de las variables de proceso que deben medirse para lograr el grado de conocimiento requerido sobre la planta. Este problema lo abordamos mediante el uso de la metaheurística *Simulated Annealing* hibridizada con estrategias especialmente diseñadas para el problema tales como Oscilación Estratégica. Además, se están incorporando métodos de simulación que permiten evaluar el diseño bajo un amplio abanico de situaciones.

**Palabras claves:** Metaheurísticas, Optimización, Diseño de red de sensores, Simulated Annealing.

## Contexto

Esta línea de investigación se desarrolla en el marco de dos proyectos de investigación, acreditados en distintas universidades argentinas. Uno llevado a cabo en el Grupo de Optimización (GOp), de la Facultad de Ingeniería de la Universidad Nacional de Río Cuarto. El segundo se desarrolla en el Laboratorio de Investigación de Sistemas Inteligentes (LISI), de la Facultad de Ingeniería de la Universidad Nacional de La Pampa, acreditado por dicha facultad y dirigido por la Dra. Minetti.

El monitoreo en una planta química es absolutamente crucial y tiene considerables impactos

en aspectos económicos, de seguridad o de control. Este monitoreo también debe garantizar el conocimiento sobre el estado del proceso en cualquier momento. La información es recopilada por sensores distribuidos en toda la planta, responsables de medir y transmitir los valores de magnitudes como temperatura, humedad, presión, etc. El conjunto de dispositivos utilizados en la medición se denomina red de sensores (*Sensor Network*, SN). El diseño de una SN se realiza sistemáticamente formulando un problema de optimización llamado problema de diseño de la red de sensores (*Sensor Network Design Problem*, SNDP), que es un problema de optimización discreta. En escenarios de trabajo reales, el número de variables involucradas es bastante grande y la formulación de SNDP puede ser más o menos compleja dependiendo de los criterios de desempeño y las restricciones que se le imponen.

SNDP fue formulado por Bagajewicz [1] y resuelto por Nguyen [2], usando un nuevo método de búsqueda de árbol. Este método explota ciertas propiedades de costo de los diferentes nodos en el árbol para podar eficientemente nodos no óptimos. Otros enfoques modelaron el problema combinando la programación entera y no lineal, resolviéndolo por medio de búsquedas de árbol que utilizan estrategias de inspección tanto en profundidad como en amplitud. La principal desventaja de todos estos métodos es la excesiva demanda de tiempo de cómputo, así como la poca flexibilidad ante cambios en la formulación del problema. En consecuencia, para lograr una optimización eficiente es importante contar con una herramienta

que permita resolver el problema para diferentes formulaciones de diferente complejidad y tamaño, en períodos de tiempo razonables.

Los diferentes métodos metaheurísticos mitigan la desventaja mencionada anteriormente porque resuelven eficientemente casos de problemas mono o multiobjetivos. En este sentido, la literatura presenta algoritmos genéticos [3], inteligencia basada en cúmulo de partículas o enjambres [4], entre otras metaheurísticas basadas en la población. Además, en la literatura se encuentran metaheurísticas híbridadas que resuelven este problema. En una de las más recientes, Carnero et al. [5] proponen la técnica PBIL\_SOTS, que combina el algoritmo de estimación de distribuciones con el algoritmo *Tabu Search* (TS) mejorado mediante el uso de una oscilación estratégica (SOTS).

En contraste con las metaheurísticas basadas en población ya mencionadas, el algoritmo *Simulated Annealing* (SA) [6] puede clasificarse en el grupo basado en trayectoria y ha demostrado ser un método eficiente para resolver muchos problemas de optimización combinatoria dura [7]. En esta línea de investigación, presentamos un algoritmo híbrido basado en *Simulated Annealing* (HSA) como metaheurística principal y *Tabu Search* integrada con una oscilación estratégica (SOTS), como metaheurística subordinada, para resolver el SNDP. Su adaptación a la resolución de un problema particular implica, entre otros aspectos, hacer un ajuste de sus parámetros que puedan obtener el mejor rendimiento de la técnica propuesta. Por lo tanto, el trabajo también se enfoca en el ajuste de los parámetros de control algorítmico para alcanzar un equilibrio entre la calidad de la solución y el consumo de tiempo. El algoritmo de optimización requiere utilizar en todos los casos un método de reconciliación de datos y verificación de cumplimiento de las restricciones impuestas. Dependiendo de la naturaleza de éstas últimas, el método de verificación puede ser analítico o bien llevado a cabo a través de técnicas de simulación. En particular, si se imponen restricciones sobre la confiabilidad asociada a un conjunto de variables claves, es posible evaluarlos mediante la implementación de métodos de Montecarlo para su evaluación.

## Desarrollo

En esta sección se describe la línea de investigación ya mencionada, introduciendo una explicación más detallada del problema y del algoritmo propuesto para resolverlo.

### Problema de diseño de la red de sensores (SNDP)

El SNDP se puede definir como el problema de encontrar una estructura de sensores que satisfaga un conjunto de restricciones a la vez que se optimiza un criterio de desempeño determinado. Una formulación particular en la que se ha trabajado es aquella donde se pretende diseñar una red de instrumentación de costo mínimo que satisfaga restricciones de precisión y estimabilidad sobre un conjunto de variables claves. Formalmente esto puede ser expresado tal como se muestra en la Ecuación (1), donde  $\mathbf{q}$  es un vector  $n$ -dimensional de variables binarias tal que  $q_i = 1$  si se mide la variable  $i$  y  $q_i = 0$  en caso contrario,  $\mathbf{c}^T$  es el vector de costo;  $\hat{\sigma}_k$  es el desvío estandar de la variable  $k$  contenida in  $S_\sigma$  después de aplicar un procedimiento de conciliación de datos [8], y  $E_l$  representa el grado de estimabilidad de la variable  $l$ -ésima incluida en  $S_E$ . Además,  $S_\sigma$  y  $S_E$  son el conjunto de variables clave del proceso con requisitos de precisión y capacidad para estimarse, respectivamente.

$$\min \mathbf{c}^T \mathbf{q} \quad (1)$$

sujeto a:

$$\begin{aligned} \hat{\sigma}_k(\mathbf{q}) &\leq \sigma_k^*(\mathbf{q}) & \forall k \in S_\sigma \\ E_l &\geq 1 & \forall l \in S_E \\ \mathbf{q} &\in \{0, 1\}^n \end{aligned}$$

En esta formulación, las mediciones están sujetas a errores aleatorios no correlacionados; solo hay un dispositivo de medición potencial para cada variable, y no hay restricciones para la localización de instrumentos. La viabilidad de las restricciones se puede verificar ejecutando una clasificación variable y un procedimiento de reconciliación de datos.

## SA híbrido para resolver SNDP

SA funciona como metaheurística principal con una búsqueda local ad hoc subordinada, inspirada en la búsqueda tabú con la técnica de oscilación estratégica, SOTS, dando lugar al HSA (*Hybrid Simulated Annealing* [9]). Esta hibridación se aplica en dos niveles: en el primero, para generar una solución inicial, y en el segundo, para mejorar la solución durante el proceso de enfriado.

El esquema de perturbación de la solución actual se lleva a cabo para generar una solución candidata a partir de la actual. Ésta consiste en realizar una cierta cantidad de intercambios, que van desde variables medidas a otras no medidas y viceversa. Esta mutación se aplica sobre cada variable con una cierta probabilidad (llamada Pswap) [10]. Además, la temperatura se actualiza al aplicar el esquema de planificación geométrico.

La primera cuestión de diseño en HSA se relaciona con los componentes de búsqueda que se fijan en función del SNDP, es decir, definir la representación de una solución. En este caso, una solución al problema SNDP es un vector  $n$ -dimensional  $\mathbf{q} = \{q_1, q_2, \dots, q_n\}$  de variables binarias, como se especifica en la definición del problema. Para evaluar una solución se usa la ecuación (2) como función objetivo, donde  $\sum_{i=1}^n c_i$  es el costo de la SN cuando se miden todas las variables (límite superior de la función objetivo del SNDP). Cuando no se cumplen las restricciones, la ecuación (3) calcula  $S(\mathbf{q})$ , donde  $rno$  y  $rnp$  representan el número de restricciones de observabilidad y precisión que no se cumplen para un vector solución.

$$H = \begin{cases} \sum_{i=1}^n c_i q_i & \text{si } \mathbf{q} \text{ es factible} \\ \sum_{i=1}^n c_i (1 + S(\mathbf{q})) & \text{si } \mathbf{q} \text{ no es factible} \end{cases} \quad (2)$$

$$S(\mathbf{q}) = \frac{rno}{S_E} + \frac{1}{rnp} \sum_{i=1}^r np \frac{\hat{\sigma}_i - \sigma_i^*}{\hat{\sigma}_i} \quad (3)$$

El segundo punto de diseño involucra los principales componentes del HSA que son variables durante el proceso de búsqueda: temperatura inicial, esquema de enfriamiento y longitud

de la cadena de Markov. A fin de estudiar el impacto de diferentes temperaturas iniciales en el rendimiento de HSA, utilizamos valores iniciales,  $T_s$ , pertenecientes a  $\{1, 900\}$ . A la hora de analizar el efecto producido por el esquema de enfriamiento, proponemos cuatro versiones de HSA:  $HSA_{Prop}$  adopta una planificación proporcional [6],  $HSA_{Exp}$  usa el esquema exponencial [6],  $HSA_{Log}$  emplea la opción logarítmica [11] y, por último,  $HSA_{Rand}$  que aplica un esquema aleatorio [12]. Finalmente, también consideramos tres formas diferentes de calcular la longitud de la cadena de Markov (MCL). Al principio, MCL se calcula estáticamente y se mantiene constante durante 30 iteraciones (número comúnmente utilizado en la literatura), denominado  $MCL_s$ . Las otras dos formas implementan las técnicas adaptativas propuestas por Cardoso et al. [13] y Ali et al. [14], denominadas  $MCL_{a1}$  y  $MCL_{a2}$ , respectivamente. Al combinar todos los enfoques antes mencionados, se obtienen un total de 24 variantes de HSA para resolver el SNDP [15].

## Resultados Obtenidos

Con el propósito de evaluar el rendimiento de HSA para resolver SNDP, se han considerado un conjunto de 5 casos de estudio que comprenden procesos de diferente complejidad y tamaño, cuya operación puede ser representada tanto por modelos lineales como no lineales [16], [17]. La complejidad del conjunto de restricciones impuestas a todos los casos de estudio se encuentra en [5].

La efectividad de las variantes de HSA se evalúa, empíricamente y con soporte estadístico, en términos de calidad de la solución, tiempo de ejecución y cantidad de evaluaciones [15]. A partir de esta evaluación se concluye que las variantes de HSA, que calculan estáticamente la longitud de la cadena de Markov (MCL), siempre obtienen la mejor solución con el mínimo esfuerzo para cuatro casos. En cambio, para el caso restante, el más complejo, se logra una compensación entre calidad y tiempo cuando HSA establece la temperatura inicial en un valor alto ( $T_s = 900$ ), usa un MCL adaptativo ( $MCL_{a1}$ ) y aplica el esquema de enfriamiento proporcional. Todo lo anterior permite

inferir que HSA es un algoritmo competitivo para resolver casos realistas de SNDP.

## Formación de Recursos Humanos

Cada año se incorporan a los proyectos alumnos avanzados en la carrera Ingeniería en Sistemas, quienes trabajan en temas relacionados a la resolución de problemas de optimización usando técnicas inteligentes, con el objeto de guiarlos en el desarrollo de sus tesis de grado y, también, de formar futuros investigadores científicos. Por otra parte, los docentes-investigadores que integran los proyectos realizan diversos cursos de posgrado relacionados con la temática del proyecto, con el objetivo de sumar los créditos necesarios para cursar carreras de posgrado.

### REFERENCES

- [1] M. Bagajewicz, "Design and retrofit of sensor networks in process plants," *AIChE Journal*, vol. 43, no. 9, pp. 2300–2306, 1997.
- [2] D. Nguyen and M. Bagajewicz, "New efficient breadth-first/level traversal tree search method for the design and upgrade of sensor networks," *AIChE Journal*, vol. 57, no. 5, pp. 1302–1309, 2011. [Online]. Available: <https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.12332>
- [3] C. Gerkens and G. Heyen, "Use of parallel computers in rational design of redundant sensor networks," *Computers and Chemical Engineering*, vol. 29, no. 6 SPEC. ISS., pp. 1379–1387, 2005.
- [4] Y.-J. He and Z.-F. Ma, "Optimal design of linear sensor networks for process plants: A multi-objective ant colony optimization approach," *Chemometrics and Intelligent Laboratory Systems*, vol. 135, pp. 37–47, 2014.
- [5] M. Carnero, J. L. Hernández, and M. Sánchez, "Optimal sensor location in chemical plants using the estimation of distribution algorithms," *Industrial & Engineering Chemistry Research*, vol. 57, no. 36, pp. 12 149–12 164, 2018.
- [6] S. Kirkpatrick, C. G. Jr, and M. Vecchi, "Optimization by simulated annealing," *Science*, no. 220, pp. 671–680, 1983.
- [7] E. Talbi, *Metaheuristics: From Design to Implementation*. Wiley Publishing, 2009.
- [8] M. Bagajewicz and M. Sánchez, "Reallocation and upgrade of instrumentation in process plants," *Computers Chemical Engineering*, vol. 24, no. 8, pp. 1945 – 1959, 2000.
- [9] J. Hernandez, C. Salto, G. Minetti, M. Carnero, and M. Sanchez, "Hybrid simulated annealing for optimal cost instrumentation in chemical plants," *Chemical Engineering Transactions*, vol. 74, pp. 709–714, May 2019.
- [10] K.-L. Du and M. N. S. Swamy, *Search and Optimization by Metaheuristics: Techniques and Algorithms Inspired by Nature*, 1st ed. Birkhäuser Basel, 2016.
- [11] B. Hajek, "Cooling schedules for optimal annealing," *Mathematics of Operations Research*, vol. 13, no. 2, pp. 311–329, 1988.
- [12] C. Bermudez, C. Salto, and G. Minetti, "Solving the multi-period water distribution network design problem with a hybrid simulated annealing," in *Computer Science – CACIC 2018*, P. Pesado and C. Aciti, Eds. Cham: Springer International Publishing, 2019, pp. 3–16.
- [13] M. F. Cardoso, R. L. Salcedo, and S. F. de Azevedo, "Nonequilibrium simulated annealing: A faster approach to combinatorial minimization," *Ind. Eng. Chem. Res.*, vol. 33, pp. 1908–1918, 1994.
- [14] M. Ali, A. Törn, and S. Viitanen, "A direct search variant of the simulated annealing algorithm for optimization involving continuous variables," *Computers & Operations Research*, vol. 29, no. 1, pp. 87 – 102, 2002.
- [15] J. Hernandez, C. Salto, G. Minetti, M. Carnero, C. Bermudez, and M. Sanchez, "Optimal instrumentation: Adjustment and hybridization of a simulated annealing based technique," in *XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019)*, Oct. 2019, pp. –.
- [16] M. Bhushan and R. Rengaswamy, "Design of sensor network based on the signed directed graph of the process for efficient fault diagnosis," *Industrial & Engineering Chemistry Research*, vol. 39, no. 4, pp. 999–1019, 2000. [Online]. Available: <https://doi.org/10.1021/ie990383d>
- [17] H. Smith and N. Ichiyen, "Computer adjustment of metallurgical balances," *Computer Applications and Process Control (CIM) Bulletin*, vol. 66, pp. 97–100, 1973.