

Enfoque de Monte Carlo aplicado a la estimación de variables en el diseño de la red de sensores en plantas químicas

Carlos Bermudez¹, José Luis Hernandez², Mercedes Carnero²,
Hugo Alfonso¹, Gabriela Minetti¹, Carolina Salto¹

¹Laboratorio de Investigación en Sistemas Inteligentes (LISI)
Facultad de Ingeniería - Universidad Nacional de La Pampa

e-mail: ¹{bermudezc, alfonsoh, minettig, saltoc,@ing.unlpam.edu.ar}

²Grupo de Optimización (GOp)

Facultad de Ingeniería - Universidad Nacional de Río Cuarto

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

Resumen

Conocer el estado real de una planta química tiene un impacto significativo en su eficiencia energética y en el cumplimiento de las normas medioambientales y de seguridad. Este conocimiento se puede obtener instalando una red de sensores adecuada en la fase de diseño de la planta y aplicando procedimientos de tratamiento de datos durante su operación. De esta manera, los datos proporcionados por los sensores se transforman en información consistente para ser utilizada como insumo para una metodología de monitoreo, control y optimización en línea. La cantidad de sensores a instalar en la planta depende del conjunto de variables del proceso que se deben conocer de manera confiable para ejecutar los procedimientos antes mencionados. La selección óptima de instrumentación de plantas químicas se define como el problema de diseño de redes de sensores (SNDP).

Esta línea de investigación considera la propuesta del SNDP para el monitoreo del desempeño de procesos. El funcionamiento de la planta está representado por un sistema no lineal de ecuaciones algebraicas. El diseño minimiza el costo de instrumentación y también satisface la observabilidad de un conjunto requerido de variables de proceso y la precisión de algunas de ellas. La precisión se evalúa en términos de la desviación estándar de las estimaciones de variables calculadas utilizando el enfoque de Monte Carlo.

Palabras claves: Optimización, Red de sensores, simulación, algoritmo de Monte Carlo

CONTEXTO

Un conocimiento fiable y completo del estado actual de una planta química es esencial para su seguimiento, control regulatorio y de supervisión, optimización en tiempo real, planificación y programación, etc. ([1]). La calidad y disponibilidad de esa información dependen en gran medida de la red de sensores (SN) instalada en el proceso.

La cantidad de sensores a instalar en la planta depende del conjunto de variables del proceso que se deben conocer de manera confiable para ejecutar los procedimientos antes mencionados. La selección óptima de instrumentación de plantas químicas se define como el problema de diseño de redes de sensores (SNDP)([2], [3]). Debido a que la estimación de las variables requeridas se puede lograr ubicando diversos tipos de sensores para medir diferentes variables del proceso, surge un enorme problema de optimización combinatoria incluso para plantas de pequeña escala. Su solución determina el conjunto de variables del proceso a medir que optimiza los criterios seleccionados y cumple un conjunto de restricciones ([4], [5], [6], [7], [8]). En general, el número de variables que intervienen en estos problemas para un escenario de trabajo real es bastante grande y la formulación puede ser más o menos compleja dependiendo

de los criterios de actuación y del conjunto de restricciones utilizadas.

El objetivo de esta línea de trabajo es proporcionar una herramienta para evaluar la precisión de las estimaciones de variables del proceso en el contexto del diseño óptimo de una red de sensores en plantas químicas. Una de las posibles formulaciones para el diseño óptimo de un sistema de instrumentación para tareas de monitoreo es la solución de problemas de optimización no lineal con restricciones, donde la función objetivo es el costo del instrumento y las restricciones son la observabilidad y precisión global asociada a la colocación de un sensor. Cuando se utiliza un enfoque metaheurístico para resolver este problema [9], [10], [11], se necesita una metodología para calcular las restricciones para evaluar la calidad de una solución propuesta. Se ha seleccionado una técnica de simulación para resolver la precisión asociada a un conjunto de medidas basada en el algoritmo de Monte Carlo. El simulador requiere de una metodología de clasificación de variables y una función de conciliación de datos que consiste en resolver otra optimización no lineal.

El desarrollo de esta línea de investigación parte de un trabajo colaborativo enmarcado en dos proyectos de investigación acreditados en distintas universidades argentinas. Uno de ellos es llevado a cabo por el Grupo de Optimización (GOp) perteneciente a la Facultad de Ingeniería de la Universidad Nacional de Río Cuarto, en tanto que 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.

1. DESARROLLO

Esta sección presenta una explicación detallada del problema SNDP y del algoritmo de Monte Carlo utilizado para resolver la precisión del conjunto de variables medidas.

A. Problema de diseño de la red de sensores

Supongamos que \mathbf{R} es un conjunto de ecuaciones algebraicas no lineales que representa la operación de un proceso en condiciones de estado estacionario (see Equation 1).

$$\mathbf{R}(\mathbf{z}) = \mathbf{R}(\mathbf{x}, \mathbf{u}) = 0 \quad (1)$$

donde \mathbf{z} es el vector de n dimensiones de las variables del proceso mientras que \mathbf{x} y \mathbf{u} son los vectores de las variables medidas y no medidas, respectivamente. La selección óptima de instrumentos para el monitoreo de procesos consiste en determinar la partición del vector \mathbf{z} en los vectores \mathbf{x} y \mathbf{u} que optimice un criterio dado y satisfaga un conjunto de restricciones. En general, se imponen restricciones al grado de estimabilidad de las variables clave del proceso y también a la calidad de sus estimaciones. Un problema de selección particular obtiene la red de sensores de costo mínimo que satisface un grado de estimabilidad mayor o igual a 1 y considera restricciones de precisión para un conjunto dado de variables requeridas. Este problema está planteado por la Ecuación 2:

$$\min \mathbf{c}^T \mathbf{q} \quad (2)$$

sujeto a:

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

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 estándar de la variable k contenida en S_σ después de aplicar un procedimiento de conciliación de datos ([12]) y E_l representa el grado de estimabilidad de la variable l -ésima incluida en S_E . Además, S_σ y S_E son el conjunto de variables clave del proceso con requisitos de precisión y capacidad para estimarse, respectivamente.

B. Simulación por Monte Carlo

En esta formulación, las mediciones están sujetas a errores aleatorios no correlacionados; solo existe un dispositivo de medición potencial para cada variable y no existen restricciones para la localización de instrumentos. Además, la viabilidad de las restricciones se puede comprobar ejecutando procedimientos de clasificación de variables y conciliación de datos [13].

La restricción de precisión asociada con la estimación de una variable requerida se puede evaluar explícitamente en sistemas lineales, pero

no en sistemas no lineales. Un enfoque consiste en realizar una linealización de las ecuaciones de equilibrio que gobiernan el modelo de planta alrededor de un punto operativo conocido. Luego se puede calcular la precisión de las variables requeridas utilizando una metodología de conciliación de datos lineal. En este caso, es posible tener una expresión explícita para el cálculo de la desviación con expresiones que impliquen calcular la inversa de matrices que pueden tener números de condición altos [3]. Una segunda opción es implementar un método de Monte Carlo (MC), que tiene la ventaja de ser flexible y simple, pero cuya precisión depende del tamaño de la muestra con la que se realiza la simulación.

En este caso, se usa una metodología basada en la segunda opción para una estimación de precisión que se pueda implementar en la metodología de colocación óptima de sensores. La viabilidad de las soluciones determinadas mediante un procedimiento de simulación implica la resolución de problemas de optimización no lineal en un segundo nivel.

La Simulación Monte Carlo es un procedimiento numérico estocástico ampliamente utilizado para simular el comportamiento de un fenómeno o un proceso real que involucra variables aleatorias de manera que se conocen sus funciones de densidad de probabilidad. En este problema, el muestreo del vector de medidas aleatorias se realiza considerando la función de densidad de probabilidad de cada medida. En tanto que, la convergencia de este algoritmo de simulación garantiza que el valor de desviación estimado esté limitado dentro de un pequeño rango de tolerancia.

2. RESULTADOS OBTENIDOS

El caso de estudio es un reactor de tanque agitado continuo (CSTR) [14], cuyo modelo comprende 13 variables (caudales totales, composiciones y temperaturas) con 5 balances de masa y energía.

Para este caso de estudio, la convergencia de la media muestral de los valores conciliados con los verdaderos se logra más rápido que la estabilización de la desviación estándar muestral alrededor de un cierto valor. Además, el número de pruebas de simulación de MC necesarias para lograr la convergencia de la desviación estándar de la muestra depende de la solución analizada. El estudio de

simulación realizada permite inferir que el método MC requiere $N=1000$ ensayos, en promedio, para proporcionar una estimación confiable de la precisión de una variable requerida. Por lo tanto, un límite superior del esfuerzo computacional total necesario para verificar la viabilidad de las restricciones de precisión para una solución potencial es $N \#$ variables requeridas. Para el problema CSTR, el número máximo de llamadas a la función de reconciliación de datos para cada solución a evaluar es 3000. Para resolver problemas de mayor tamaño, se debe revisar la utilización de MC [15].

3. FORMACIÓN DE RECURSOS HUMANOS

Cada año se incorporan a los proyectos alumnos avanzados en la carrera Ingeniería en Sistemas de la UNLPam, 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] P. Christofides, J. Davis, N. El-Farra, D. Clark, H. K., and J. Gipsen, "Smart plant operations: Vision, progress and challenges." *AIChE J*, vol. 53, pp. 2734–2741, 2007.
- [2] M. Bagajewicz, "Design and retrofit of sensor networks in process plants," *AIChE Journal*, vol. 43, no. 9, pp. 2300–2306, 1997.
- [3] 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.
- [4] 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.
- [5] P. Kotecha, M. Bhushan, R. Gudi, and M. Keshari, "A duality based framework for integrating reliability and precision for sensor network design," *Journal of Process Control*, vol. 18, no. 2, pp. 189–201, 2008.
- [6] P. Sen, K. Sen, and U. M. Diwekar, "A multi-objective optimization approach to optimal sensor location problem in IGCC power plants," *Applied Energy*, vol. 181, pp. 527–539, 2016.
- [7] P. Paul, D. Bhattacharyya, R. Turton, and S. E. Zitney, "Sensor network design for maximizing process efficiency: an algorithm and its application," *AIChE Journal*, vol. 61, no. 2, pp. 464–476, 2015.

- [8] G. Sambito, M. and Freni, “Strategies for improving optimal positioning of quality sensors in urban drainage systems for non-conservative contaminants.” *Water.*, vol. 13, pp. 1–14, 2021.
- [9] J. L. Hernández, C. Salto, G. F. Minetti, M. Carnero, C. Bermúdez, and M. Sánchez, “Tuning a hybrid sa based algorithm applied to optimal sensor network design,” *Journal of Computer Science & Technology*, vol. 20, no. 1, 2020.
- [10] 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. –.
- [11] 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.
- [12] M. Bagajewicz and M. Sanchez, “Reallocation and upgrade of instrumentation in process plants,” *Computers & Chemical Engineering*, vol. 24, no. 8, pp. 1945 – 1959, 2000.
- [13] J. Romagnoli and M. Sánchez, *Data Processing and Reconciliation for Chemical Process Operations*. Academic Press: San Diego, 1999.
- [14] 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.
- [15] J. Hernández, S. Simón, M. Carnero, G. Minetti, and C. Salto, “On the precision evaluation in non-linear sensor network design,” in *Libro de Actas: XXIX Congreso Argentino de Ciencias de la Computación - CACIC 2023*, 2023, p. 48.