Aportaciones Matemáticas
Comunicaciones 17 (1996) 25–34.

Optimización del diseño de una red de distribución de agua potable[1]

M. LENTINI [2] E. LOYOLA [3] F. MEMBRILLO [4]
J.L. MORALES [5] R. PENICHE[6] Z. RAMOS[7] J. RAMÍREZ [8]
D. ROMERO [9] G. SANTANA [10] V. TZACHKOV[11]

Agosto de 1995
(versión revisada)

[1]Problema presentado por el Instituto Mexicano de Tecnología del Agua en el GRUPO DE ESTUDIO CON LA INDUSTRIA, organizado por la *Sociedad Matemática Mexicana* en colaboración con el *European Consortium for Mathematics in Industry.* Oaxaca, Oax., febrero 1995

[2]Universidad Simón Bolívar
[3]Instituto Nacional de Estadística, Geografía e Informática
[4]Instituto de Matemáticas, UNAM
[5]Instituto Tecnológico Autónomo de México
[6]Universidad Autónoma de Querétaro
[7]Facultad de Ciencias, UNAM
[8]Instituto de Matemáticas, UNAM
[9]Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas, UNAM
[10]Centro de Investigación y de Estudios Avanzados, IPN
[11]Instituto Mexicano de Tecnología del Agua

Resumen

En el presente reporte se presentan los resultados obtenidos por el grupo de trabajo que estudió el problema de diseñar de manera óptima, una red de distribución de agua potable. Esencialmente se discuten dos clases de estrategias. En primer lugar, aquellas cuya finalidad es reducir significativamente los recursos computacionales requeridos por los algoritmos implementados por el IMTA. Estos algoritmos son de caracter heurístico y generan una solución factible que no es óptima. En ciertos casos se sabe que las soluciones obtenidas por dichos algoritmos están relativamente lejos del óptimo y no son aceptables desde el punto de vista del diseñador. La segunda clase de estrategias propuestas, está destinada precisamente a aliviar este problema. Se sugieren técnicas originadas en optimización continua y en flujo en redes.

Por diferentes razones es necesario realizar grandes inversiones en la proyección, extensión y rehabilitación de redes de distribución de agua destinada al consumo humano. En consecuencia, es de vital importancia utilizar métodos eficientes para generar diseños de costo mínimo que a la vez que satisfagan un cierto conjunto de especificaciones técnicas.

El problema a resolver consiste en asignar diámetro a cada tramo de una red de agua potable con m tramos de tubería y n nodos de consumo y abastecimiento, de tal manera que el costo total de la red resulte mínimo. Se supone que la red tiene una topología preestablecida y que la función de costo depende del diámetro y de la longitud de los tubos utilizados. Por otra parte, el diseñador solamente dispone de un número limitado de diámetros accesible en el mercado. Finalmente, las especificaciones del diseño requieren que la presión de entrega en los nodos de consumo sea superior a un cierto valor límite y que la velocidad del fluído en cada tramo de la tubería no rebase cierta tolerancia establecida por el fabricante.

En la literatura especializada el problema anterior está ampliamente documentado y se le ha reportado como una tarea extremadamente compleja, ver por ejemplo [9]. Dado el astronómico número de asignaciones posibles y las restricciones impuestas por el diseño, la enumeración exhaustiva no es un método recomendable en casos de aplicaciones concretas. Por otra parte, desde el punto de vista computacional, el problema ha sido clasificado como $\mathcal{NP}$ completo; lo cual significa en términos prácticos, que el número de operaciones requerido para obtener la solución exacta es probablemente de orden exponencial. Por estas razones, diversos autores han propuesto algoritmos que obtienen solamente una aproximación a la solución. En general, los métodos descritos en la literatura pueden ser clasificados de acuerdo al siguiente esquema

- Técnicas heurísticas que obtienen una solución factible.
- Técnicas basadas en optimización continua que resuelven casos relativamente pequeños con un nivel aceptable de cercanía al valor óptimo.

Una vez presentado el problema, el grupo de trabajo se concentró en el estudio de los algoritmos implantados en el IMTA. Estos algoritmos pueden ser clasificados dentro de la primera categoría, es decir obtienen un conjunto de valores para los diámetros que cumple con la especificaciones de diseño. Sin embargo, por experiencia se sabe que en ciertos casos la aproximación obtenida está relativamente lejos del óptimo, careciendo en consecuencia de valor práctico. Por lo tanto la primera meta del grupo de estudio consistió en obtener mejoras en los algoritmos existentes. Posteriormente, el grupo propuso y discutió algunas estrategias pertenecientes a la segunda categoría.

El reporte está organizado de la siguiente manera: en la sección 2 formulamos el problema desde un punto de vista ligeramente diferente al presentado en [9], pero que creemos contribuye a la concepción de nuevos métodos de resolución. En la sección 3 discutimos las estrategias para mejorar los algoritmos existentes, así como su implementación computacional. En la sección 4 discutimos dos propuestas. Una de ellas basada en programación cuadrática recursiva y la otra en flujo en redes.

Antes de iniciar la discusión, introduciremos alguna notación que será de utilidad durante el resto del documento. Los conjuntos serán denotados con letras caligráficas mayúsculas. Todos los vectores serán vectores columna, la transpuesta de una matriz A será denotada por A^T. El símbolo $||\cdot||$ denotará a la norma-2. Finalmente, dado un vector x, el símbolo X denotará a la matriz diagonal cuyas entradas son las componentes de x.

2 Formulación del problema

Matemáticamente, la topología de una red de m tramos y n nodos se puede representar por medio de una matriz A de $m \times n$, en la cual la entrada $a_{i,j}$ está definida como sigue

$$a_{i,j} = \begin{cases} \pm 1 & \text{si existe conexión entre } i, j \\ 0 & \text{en otros casos,} \end{cases}$$

en donde el signo se escoge dependiendo de la dirección del flujo entre los nodos i, j. Las especificaciones de diseño así como el cumplimiento de ciertas leyes físicas, se pueden agrupar en un conjunto de restricciones que se imponen a una función de costos que depende esencialmente de los diámetros de los tubos. Por lo tanto el problema se puede formular como el siguiente problema de optimización mixta entera

$$\begin{array}{lllll} \text{minimizar} & c(d) & & & \\ \text{sujeta a} & A^T q - q_0 & = & 0 & \\ & k(d_i)|q_i|q_i - (p_{j_{i1}} - p_{j_{i2}}) & = & 0, & i = 1, 2, \ldots, m \\ & q_i d_i^{-2} - \alpha v_i & = & 0, & i = 1, 2, \ldots, m \\ & p \geq p^{min}, \quad v \leq v^{max}, \quad d_i \in \mathcal{D} & & & \end{array} \qquad (1)$$

en donde c es una función, no necesariamente lineal, de d el vector de diámetros, $d \in \mathbf{R}^n$. El primer bloque de restricciones representa el balance de gastos en cada nodo; $q_0 \in \mathbf{R}^n$ es el vector de gastos (conocido) y $q \in \mathbf{R}^m$ es el vector de descargas (desconocidas). El segundo bloque de restricciones representa la caída de presión en cada tramo de la red, la cual depende de la descarga en los tramos y del diámetro del tubo (a través de la función empírica $k(d_i)$). El último bloque establece la dependencia entre velocidad (desconocida) y descarga (desconocida) en cada tramo de la red. Finalmente, se imponen restricciones sobre las presiones y las velocidades, así como el hecho de que los valores que pueden tomar los diámetros se encuentran discretizados, es decir $\mathcal{D} = \{d^1, d^2, \ldots, d^r\}$.

3 Modificaciones propuestas

Entre algunas de las dificultades reportadas para la resolución de (1), la falta de un punto inicial factible razonablemente bueno, ha motivado el desarrollo de técnicas que parten de una solución infactible pero de costo mínimo (máximo) y progresivamente incrementan (decrementan) los diámetros en aquellos tramos en los cuales ocurren violaciones severas a las restricciones. El método termina cuando se alcanza factibilidad para algún juego de diámetros. Típicamente, un método de esta clase se puede esquematizar de acuerdo al siguiente

Algoritmo básico

Escoger un punto inicial $d^{(0)}$ de costo mínimo

Desde $j = 0$ **repetir**

Verificar factibilidad

Si $d_i^{(j)} = d^k$ infactible **entonces**

$d_i^{(j)} \rightarrow d^{k+1}$

$j \leftarrow j + 1$

En caso contrario FIN

fin.

En términos computacionales, el algoritmo anterior es extremadamente costoso en la fase de verificación. En general, esta fase implica la resolución de un sistema de ecuaciones no lineales con algún método iterativo como el algoritmo de Newton. La segunda fase, actualiza los valores de los diámetros de acuerdo a algún criterio basado en la restricción más violada. De acuerdo a la referencia [9], el esquema utilizado consiste en

1. Identificar aquellos tramos con exceso de velocidad. Aumentar el diámetro del tramo en el cual la violación es mayor

2. Verificar factibilidad, en caso de violación, regresar al punto anterior

3. Identificar los nodos con déficit de presión. Aumentar el diámetro del tramo con mayor influencia sobre el déficit de presión

4. Verificar factibilidad, en caso de violación, regresar al punto anterior.

En el esquema anterior es claro que la fase de factibilidad puede llegar a ser invocada un gran número de veces durante cada iteración del algoritmo. Por lo tanto las modificaciones propuestas están enfocadas en las siguientes direcciones:

- Uso de estrategias de actualización de los diámetros que consideren un ascenso mínimo en la función objetivo

- Uso de métodos eficientes para resolver los sistemas de ecuaciones no lineales involucrados en la fase de verificación de factibilidad.

En lo que resta del documento, nos concentraremos en el estudio de métodos para resolver los sistemas de ecuaciones no lineales. Estos sistemas se pueden formular en función de los valores de los diámetros, es decir

$$\begin{pmatrix} A^T q - q_0 \\ q^T K|q| - \Delta p \\ D^{-2} q - \alpha v \end{pmatrix} = 0, \quad p \geq p^{min}, \quad v \leq v^{max}, \tag{2}$$

en donde las incógnitas son q, p y v, i.e. las descargas, las presiones y las velocidades respectivamente; $|q|$ denota al vector cuyas componentes son $|q_i|$. Los algoritmos implantados en el IMTA resuelven los sistemas de ecuaciones no lineales (ENLs) anteriores utilizando el método de Newton, uno de los mejores para resolver sistemas de ENLs densos y pequeños debido a sus excelentes propiedades de convergencia cuadrática local. En su forma básica este algoritmo tiene la siguiente descripción:

Algoritmo Newton

Escoger una aproximación inicial x_0

Desde $i = 0$ hasta CONVERGENCIA

resolver para Δx

$$F'(x_i)\Delta x = -F(x_i) \tag{3}$$

actualizar x_i

$$x_{i+1} = x_i + \Delta x$$

fin

en donde F' denota a la derivada de F con respecto a x.

Cuando el algoritmo anterior se aplica a (2); entonces las dificultades de tipo computacional están cercanamente asociadas con la resolución de las llamadas ecuaciones de Newton (3), ya que éstas deben ser resueltas en cada iteración. Bajo la hipótesis de que $F(x_i)$ y $F'(x_i)$ están disponibles en memoria de acceso aleatorio, el costo computacional involucrado en computar una estimación de Δx es proporcional a n^3 operaciones en punto flotante y a n^2 espacios en memoria. Si la matriz $F'(x_i)$ es *hueca*, como ocurre en la gran mayoría de las aplicaciones prácticas, entonces el uso de técnicas apropiadas que explotan dicha propiedad disminuye significativamente el costo computacional.

Sin embargo, durante las primeras iteraciones cuando las aproximaciones están relativamente lejos de la solución, no se justifica resolver "exactamente", v.gr. a la precisión de la máquina en cuestión, el sistema (3). En [1] se propone resolver la siguiente modificación de el sistema de ecuaciones (3):

$$F'(x_i)\Delta x = -F(x_i) + r_i, \tag{4}$$

en donde r_i es denotado como el *residuo* y su norma satisface la siguiente desigualdad para alguna $\eta_i \in [0, 1)$

$$||r_i|| \leq \eta_i ||F(x_i)||. \tag{5}$$

Por lo tanto, la secuencia $\{\eta_i\}$ determina la precisión con la cual se resuelven las ecuaciones de Newton (3). En ausencia de referencia alguna al mecanismo para computar una solución de ellas, la variante resultante del algoritmo de Newton se conoce como un método de *Newton inexacto.* Cuando el método utilizado para resolverlas es de naturaleza iterativa, entonces la variante obtenida es denotada como un método de *Newton truncado.*

El análisis de los métodos obtenidos al introducir la perturbación descrita anteriormente, demuestra que la modificación propuesta en (4) preserva las propiedades de convergencia local del método de Newton. Los métodos resultantes convergen localmente a diferentes tasas dependiendo de las propiedades de la sucesión $\{\eta_i\}$.

Por lo tanto proponemos que el sistema (2) se resuelva con un método truncado, en el cual los sistemas (3) son resueltos con algoritmos iterativos precondicionados basados en subespacios de Krylov. Una descripción de como implantar computacionalmente una versión del algoritmo de Newton truncado, se encuentra en [5]. En esta referencia se muestra además una comparación con la variante "exacta". Esta comparación permite concluir que la variante truncada es aproximadamente dos veces más rápida que la variante "exacta" en los problemas de prueba utilizados.

4 Estrategias sugeridas

En esta sección presentamos dos alternativas para resolver (1). La primera de ellas es una modificación de una técnica sugerida en la literatura. Esencialmente la modificación está enfocada a acelerar la convergencia y a permitir el manejo de problemas a gran escala de manera eficiente.

La segunda alternativa propuesta considera al problema desde un enfoque radicalmente diferente. En consecuencia, es necesario reformular ligeramente (1) para hacer más fácil su presentación y discusión.

4.1 Técnicas basadas en optimización continua

Recientemente, algunos autores han sugerido técnicas basadas en optimización continua para resolver (1), ver por ejemplo Hansen *et al* [4]. Esencialmente este enfoque consiste en utilizar una variante de la llamada programación lineal secuencial en la cual, en cada iteración se formula un problema lineal. La solución de este problema proporciona una dirección de descenso para el problema original. Una vez calculada la dirección de descenso, ésta es ajustada para generar un nuevo iterando dentro de la zona factible. Los experimentos reportados en [4] sugieren que la técnica es eficiente y que solamente requiere de la resolución de unos cuantos programas lineales. Sin embargo los ejemplos utilizados son relativamente pequeños en comparación con los problemas considerados como reales. Por lo tanto es de esperarse que con problemas cuyo número de nodos está en el orden de los miles, la técnica resulte excesivamente costosa. Este es el caso en ciertos problemas de refinerías, en los cuales la técnica secuencial requiere de la resolución de miles de programas lineales.

Como alternativa sugerimos que (1) se resuelva con una técnica de convergencia más rápida, como la programación cuadrática recursiva, en la cual las direcciones de descenso requieren de la resolución de un problema cuadrático a gran escala. Un buen indicador de que la aproximación lineal es insuficiente, es el hecho de que en [4] se requiere del uso de una región de confianza para estimar el grado de ajuste del modelo lineal al problema original. Los resultados numéricos sugieren que el número de iteraciones lineales se puede reducir incrementando el tamaño de la región de confianza. Sin embargo, el costo computacional de mantener a los iterandos en la zona factible es muy alto.

Existen implementaciones computacionales de diversas variantes del método descrito anteriormente capaces de resolver problemas a gran escala (ver por ejemplo [7]), sin embargo es necesario modificarlas substancialmente para poder aplicarlas a la resolución de (1).

4.2 Técnicas basadas en flujo en redes.

En esta alternativa se considera al problema (1) desde la óptica de las técnicas de flujo en redes. Para fines de claridad reformularemos (1) de la siguiente manera:

$$\begin{array}{lllll} \text{minimizar} & \sum L_{i,j} c_{i,j} & & \\ \text{sujeta a} & p_i = H_i - z_i & \geq & p_{min} \\ & v_{i,j} = Q_{i,j}/A_{i,j} & \leq & \hat{v}_{i,j} \\ & H_i - H_j & = & KQ_{i,j}^2 \\ & \sum Q_{i,j} & = & q_i, \end{array}$$

en donde los índices i, j están referidos a cada arco de la red; $L_{i,j}$ y $c_{i,j}$ representan longitud de cada tramo y su costo correspondiente. Los parámetros conocidos son: la velocidad máxima en cada tramo $\hat{v}_{i,j}$; la demanda en cada nodo de la red q_i y la elevación de cada nodo z_i. Por otra parte, los parámetros dependientes de los diámetros son: la carga en el nodo i-ésimo H_i; la presión en el nodo $i$ésimo; la velocidad del fluído en cada arco $v_{i,j}$; el flujo en cada arco $Q_{i,j}$ y la capacidad de cada arco $A_{i,j}$.

Con la formulación anterior e ignorando las restricciones

$$\begin{array}{rcl} p_i = H_i - z_i & \geq & p_{min} \\ H_i - H_j & = & KQ_{i,j}^2, \end{array}$$

el problema se podría considerar como un problema de flujo de valor dado a costo mínimo con cotas inferiores y superiores en los arcos de la siguiente manera:

- Se fija el diámetro mayor d^r para todos los arcos de la red . Con esto queda determinada la orientación de cada uno de ellos y la cota para cada arco.

- Los nodos tienen determinada demanda q_i.
- Existen algunos nodos fuente que tienen producción de flujo f_i.
- En la red se cumple que $\sum q_i = \sum f_i$
- Se considera la elevación de cada tramo $L_{i,j}$
- Se considera la siguiente modificación de la red:

  Crear dos nodos ficticios, uno fuente F y uno sumidero S. El nodo F se conecta con cada uno de los nodos fuente, por medio de un arco con cota inferior 0, cota superior f_i y costo 0. El nodo S se conecta con cada nodo, de la red original, que tenga demanda q_i, por medio de un arco de cota inferior q_i cota superior ∞ y costo 0.

- El flujo de la red satisface la Ley de Kirchoff con lo cual se cumple $\sum Q_{i,j} = q_i$.

En esta red lo que se pretende entonces es determinar el flujo de valor $\sum q_i$ a costo mínimo, entendiendo como valores de costo las $L_{i,j}$ y considerando cotas inferiores y superiores en los arcos.

Una vez resuelto, los valores del flujo en cada arco determinarán los diámetros a considerar en cada uno de ellos, dependiendo de la relación:

$$Q_{i,j} \leq \hat{v}_{i,j} A_{i,j},$$

que corresponde a las restricciones sobre las velocidades

$$v_{i,j} = Q_{i,j}/A_{i,j} \leq \hat{v}_{i,j}.$$

Con esta formulación se garantiza costo mínimo en función de diámetros y longitudes.

Referencias

[1] R.S. DEMBO, S.C. EISENSTAT, AND T. STEIHAUG, *Inexact Newton Methods*, SIAM J. Numer. Anal., 19, pp. 400-408, 1982.

[2] S.C. EISENSTAT, H.C. ELMAN AND M.H. SCHULTZ, *Variational Iterative Methods for Nonsymmetric Systems of Linear Equations*, SIAM J. Numer. Anal., 20, pp. 1983.

[3] S.C. EISENSTAT AND H.F. WALKER, *Globally Convergent Inexact Newton Methods.* SIAM J. Opt. 4, pp 393-422, 1994.

[4] C.T. HANSEN, K. MADSEN AND H.B. NIELSEN, *Optimization of Pipe Networks*, Math. Prog., 52, pp. 45-58, 1991.

[5] J.L. MORALES PÉREZ, *Implantación de Métodos tipo Newton Truncado para Resolver Sistemas Grandes de ENLs.* Reporte Técnico. Departamento de Simulación, Instituto de Investigaciones Eléctricas. Cuernavaca, Mor., Octubre 1993.

[6] H.S. NIELSEN, *Methods for Analyzing Pipe Networks*, Journal of the Hydrulics Division, 115, pp. 139-157, 1989.

[7] R.W.H. SARGENT, D. MEI AND J.L. MORALES-PÉREZ, *A New SQP Algorithm for Large Scale Nonlinear Programming* Paper 137b. *Session on Advances in Optimization.* A.I.Ch.E. Annual Meeting, Miami, FL USA, 1992.

[8] Y. SAAD AND M.H. SCHULTZ, *GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems*, SIAM J. Sci. Stat. Comput., 7, pp. 856-859, 1986.

[9] V.G. TZACHKOV Y M.E. ALFONSO, *Diseño Optimo de Redes Hidráulicas mediante Análisis Consecutivos Direccionados.* Reporte Técnico. Instituto Mexicano de Tecnología del Agua.