



Spanning tree-based algorithm for hydraulic simulation of large-scale water supply networks

Huan-feng DUAN*¹, Guo-ping YU²

1. Department of Civil and Environmental Engineering, Hong Kong University of Science and Technology, Hong Kong, P. R. China

2. College of Environmental Science and Engineering, Tongji University, Shanghai 200092, P. R. China

Abstract: With the purpose of making calculation more efficient in practical hydraulic simulations, an improved algorithm was proposed and was applied in the practical water distribution field. This methodology was developed by expanding the traditional loop-equation theory through utilization of the advantages of the graph theory in efficiency. The utilization of the spanning tree technique from graph theory makes the proposed algorithm efficient in calculation and simple to use for computer coding. The algorithms for topological generation and practical implementations are presented in detail in this paper. Through the application to a practical urban system, the consumption of the CPU time and computation memory were decreased while the accuracy was greatly enhanced compared with the present existing methods.

Key words: large-scale networks; hydraulic simulation; graph theory; fundamental loop; spanning tree; efficiency

1 Introduction

Hydraulic simulation of water supply systems can be used to provide and access system information (pressure, flow rate, and hydraulic parameters) and to design and manage the water resources scheduling (economics, efficiency, reliability, and optimization). The continuity and energy equations (Eqs. (1) and (2)) are satisfied so as to bring about a balanced state for each hydraulic element of the hydraulic functions (Yan and Liu 2002):

$$\sum_{i=1}^m q_i + Q_j = 0 \quad (1)$$

$$H_{Fi} - H_{Ti} = h_i = s_i q_i |q_i|^{n-1} \quad (2)$$

where Q_j is the discharge of the j th node, q_i is the pipe flow rate through the i th pipeline, i is the i th pipeline connecting to j th node, m is the total pipe number connecting to j th node, H_{Fi} is the pressure head at the start node of the i th pipeline, H_{Ti} is the pressure head at the end node of the i th pipeline, h_i is the head loss of the i th pipeline, s_i is the total resistance coefficient of the i th pipeline, and n is the index ($n = 2$ in this study).

*Corresponding author (e-mail: ceduan@ust.hk)
Received Nov. 15, 2009; accepted Jan. 15, 2010

Usually, according to graph theory (Novak and Gibbons 1999; Gross and Yellen 2006), the relationship between the total number of pipelines (M), total number of nodal points (N), and total number of loops (L) in an urban supply system can be written as $M = N + L - 1$. Then, the solution of the matrix equation including all nodes and loops is the final hydraulic balance state for the whole network.

Ormsbee (2006) systematically summarized some important methods that this study used in hydraulic calculations. All of those methods, developed in recent decades, were based on the achievements of Cross (1936), who proposed the loop equation-based and node equation-based methods of hydraulic calculation in water networks. Here, for convenience in the following description, the loop equation-based Hardy-Cross method will be called the Hardy-Cross method and the latter the node equation method.

The Hardy-Cross (H-C) method solves the energy equations for each loop first with the initialization of satisfied continuity conditions at each nodal point and then adjusts these initial assumed values through iterative calculations, while, for the node equation (N-E) method, the nodal pressure grades are assumed firstly with the satisfaction of the energy equations, and the continuity is solved for each nodal point iteratively. Both methods use the iterative calculations because of the nonlinear property of the original equations. The methods developed from Cross's work (1936) mainly focus on improvements in practical implementation (e.g. hydraulic devices in the system) and numerical solution schemes (Bhave 1981; Young 1994; Ormsbee 2006).

Because of the complexity of a practical system, practical hydraulic calculation is complicated and time-consuming even with the above improved methods. Many researchers have focused on the optimized treatment and decomposition of the networks in order to make calculation more efficient for practical engineering (Whaley and Hume 1986; Duan and Yu 2006; Deuerlein 2008). Those works have demonstrated the feasibility of analysis of the practical system. However, those simplified skills in actual fact also increase the efficiency to some extent by sacrificing accuracy. On the other hand, the rapid and advanced development of informatics and current requirements demand real-time sharing of the systematic information data. As a result, as one of the core parts in water systems, the hydraulic calculation will normally be required to exhibit: (i) high efficiency and accuracy with the lowest possible CPU time consumption; (ii) low memory consumption for economical costs; and (iii) high facility for data processing for digital programs and management. These requirements are important and basic for real-time information sharing, especially in a large-scale and complicated urban water controlling system.

In this paper, a new algorithm, combined with graph theory techniques, is proposed to deal with the required efficient calculations. Some practical implementations and numerical procedures are then presented. The proposed method was applied to a practical urban water supply system in a city in southern China. Some conclusions are summarized at the end.

2 Problems

A hypothetical water supply system with simple hydraulic elements (Fig. 1) was first studied with the two aforementioned traditional methods (the Hardy-Cross method and the node equation method). This system contains 13 nodes (from 0 to 12), 18 pipes (from [1] to [18]) and 6 loops (from (I) to (VI)). It is assumed that the water source (node 0) has an outflow supply with constant pressure head and other nodal points have different water consumptions. The whole system is presumed to be flat. Moreover, for a general comparison, the initializations of pipe flow and nodal pressure grade in both methods are generated by the same data generator with satisfaction of continuity and energy rules, respectively.

Under the same conditions (systematic conditions and iteration precision), the results from the two methods are shown in Fig. 2. The convergence criterion for iterations is evaluated by the maximum error of pressure head of all nodal points. Based on overall inspection of the results (Fig. 2), the Hardy-Cross method has a lower convergent speed and takes more calculation steps to obtain the same expected accuracy compared with the node equation method.

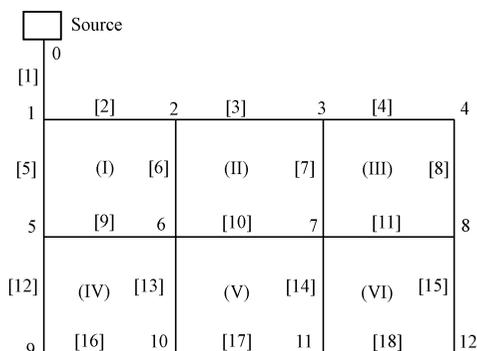


Fig. 1 Sketch of water supply system

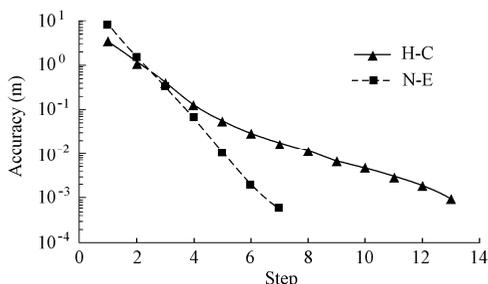


Fig. 2 Results of calculation by two methods

However, considering the efficiency (i.e., time consumption) of the process, it will take less time for each step of the Hardy-Cross method than that for the node equation method. This is because of the fact that there are fewer loops than nodal points in this case, which relates directly to the total equation number to be solved in these two schemes. This situation

is also true for the practical system. It will be investigated and discussed later in this paper.

As a result, when applying these two computation methods, especially in large-scale systems, their individual disadvantages and limitations are exposed: (1) The H-C method has a low convergence speed or accuracy, especially during later stages of the calculation; (2) the N-E method needs to deal with more data sets of equations and requires more memory, which means more economic waste. Actually, both problems are relevant to the utilization of energy resources.

Consequently, the remainder of this paper is devoted to seeking a new method of obtaining an efficient scheme, in terms of both accuracy and economics, for practical hydraulic calculation.

3 Methodology

3.1 Some concepts and techniques

3.1.1 Spanning tree

For clarity, the simple system in Fig. 1 is again adopted herein for description. Cases containing more complicated hydraulic devices (pumps, valves, multiple sources, etc.) will be discussed in the next section.

Some basic and important concepts from graph theory need to be introduced and combined with the water networks. First, the nodal point with constant pressure head is herein called the root. For example, in Fig.1, there are six loops, 13 nodal points, 18 pipelines, and one root at point 0. Taking one of the loops, loop (V) in Fig. 1, for an example, as shown in Fig. 3, there are four pipelines, which will be called the branches of the tree or branches. Hence, if one of the branches has a known value (flow rate, for example), it can be cancelled out from the loop: in this case, pipe [14]. Then this loop becomes a branched network, and the values for all other elements in this loop can be determined. The cancelled branch is therefore called the link branch or link and is expressed by the dashed line in Fig. 3.

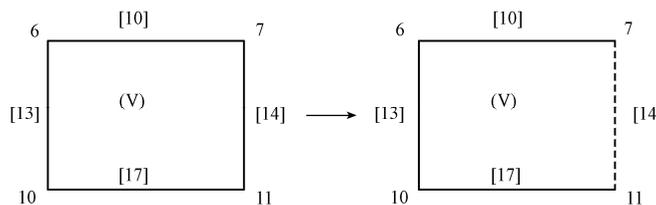


Fig. 3 Branching of loop (V)

By this definition, the system in Fig. 1, containing six loops in total, has six links in the final branching system. The decomposed network is called the spanning tree, since there is no loop in this branching system (Fig. 4). There would probably be many different spanning trees for the same system due to the different ways of decomposition. Usually, once a spanning tree is obtained, it will not be changed during future calculations. In this paper, this is called the spanning tree (S-T) method.

other hand, for fundamental loops (A) and (C), their mutual branch pipe [9] will show the contrary changing trends no matter which one of their link pipes changes. That is, if the flow rate of the link pipe [6] is increasing, the branch pipe [9] in loop (A) will also be increasing, while that in loop (C) will be decreasing. Therefore, they are defined as heterodromous, meaning that they have the contrary adjustment direction during hydraulic calculations. These definitions are useful for efficient calculation and therefore very important to the proposed computation principle later on.

3.2 Computation procedures

After the generation of the spanning trees and fundamental loops, the whole water system becomes a branched system (a tree), and it can then be calculated from the nodes and pipelines with known values to those with unknowns by the hydraulic equations (i.e., Eqs. (1) and (2)). Thus, the initialization of the link values are just like the ways for initializing the flow rate of each pipe in the H-C method or the pressure head of each nodal point in the N-E method (Yan and Liu 2002). The remainder of the work is to adjust the link values with an appropriate iteration scheme. The proposed iteration scheme for this study and the main procedures can be summarized as follows:

(1) The spanning tree and fundamental loops are generated for the original system based on the aforementioned method. To obtain an efficient search for this task by the computer program, there are many commonly used and robust algorithms in the literatures (e.g., Novak and Gibbons 1999; Gross and Yellen 2006).

(2) The flow rate value is initialized for each link pipeline. In this paper, it is suggested to be

$$q = \frac{\pi D^2}{4} \quad (3)$$

under the assumption that the velocity is equal to 1, where D is the diameter of the link pipe.

(3) The flow rate of each branch pipe from each end of the fundamental loops to the root points (such as the water source point in Fig. 1) is calculated.

(4) The pressure head of each nodal point and head loss of each pipeline from the root points to all ends of the spanning trees are computed. This path direction is just contrary to that in step (3).

(5) The difference of pressure head for the two ends of each link pipe can be calculated after step (4), and the new flow rate of that link pipe can be determined again by Eq. (1).

(6) The Jacobian matrix is created. The branch pipelines in each fundamental loop can be affected by many different link pipes from different adjacent fundamental loops. Therefore, for one fundamental loop, the energy equation should contain the influences from all of these link pipes, and the final result will be formed as a matrix equation containing L (number of fundamental loops or link pipes) equations. It is a nonlinear matrix equation with L unknown variables. An approximate linearized matrix equation can be obtained from its Taylor expansion, ignoring the high order terms:

$$A\Delta Q = \Delta H \quad (4)$$

where ΔH is the matrix of pipe head loss summation of the fundamental loop; ΔQ is the matrix of the regulated flow rate of the link pipe, which is unknown; and A is the Jacobian matrix:

$$A_{jk} = \frac{\partial H_j}{\partial Q_k} = \begin{cases} \sum_{i \in R_j} (ns_i |q_i|^{n-1}) & k = j \quad \text{contained} \\ ns_i |q_i|^{n-1} & k \neq j \quad \text{homodromous} \\ -ns_i |q_i|^{n-1} & k \neq j \quad \text{heterodromous} \\ 0 & k \neq j \quad \text{non-contained} \end{cases} \quad (5)$$

in which H_j is the closure error of pipe head loss of the j th fundamental loop; Q_k is the flow rate of k th link pipe; i is the i th tree branch pipe; R_j is the set of branch pipes in the j th fundamental loop; contained is the case of the link pipe contained in the present fundamental loop; homodromous is the case of the mutual link pipe between the present fundamental loop and another fundamental loop that has the same adjustment direction as the present fundamental loop; heterodromous is the case of the mutual link pipe between the present fundamental loop and another fundamental loop that has the contrary adjustment direction to the present fundamental loop; and non-contained is the case of the link pipe which is not contained in the present fundamental loop.

(7) The new flow rate of each link pipe can be obtained by adding the adjusted value from step (6) to the original value, then going to step (3) for iterations until the precision requirement is satisfied.

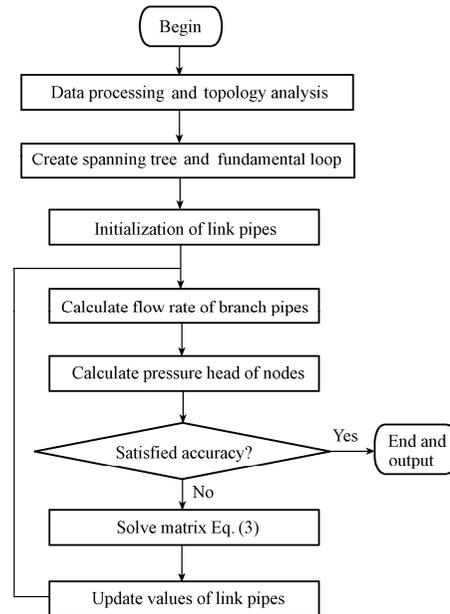


Fig. 6 Flow chart of the proposed algorithm

The numerical procedures are described in Fig. 6.

4 Practical implementations

In a practical water supply system, there are many more complicated hydraulic devices in addition to the simple pipeline and nodal points, including pump stations, valves, multiple sources, and complicated geometric pipelines. These elements are essential to a practical integrated network and therefore need to be incorporated into the proposed method.

4.1 Pump station

For a pump station, several or all working pumps can be combined as a functional pump that has the following hydraulic characteristic:

$$h_p = H_{p0} - s_p q_p^m \quad (6)$$

where h_p is the discharge head of the pump, H_{p0} is the static head of the pump, s_p is the comprehensive resistance coefficient of the pump, q_p is the discharge of the pump, and m is the index.

By involving the pump in computations, the pump is dealt with as an individual pipeline. This special pipeline has a length of 0 but has a certain resistance coefficient (s_p). Since the pump is always used to increase the energy or pressure head for this pipeline (unidirectional), the head loss for this pipeline will be $-h_p$. Therefore, Eq. (2) can be rewritten for a pump pipeline case:

$$H_{Fi} - H_{Ti} = h_i = -h_p = s_p q_p |q_p|^{m-1} - H_{p0} \quad (7)$$

The same form as in Eq. (5) can be used for the coefficient of the Jacobian matrix by substituting the index n with m .

In particular, a limiter should be added in the calculation process. That is, when the calculated pressure head difference (h_i) is positive in Eq. (7), it will be compulsively assigned a value of 0 in the iterative calculations, indicating that the pump is unidirectional.

4.2 Valves

As with the pump, valves in water distribution are treated as valve pipelines with lengths of 0, but have their own resistances. It is also assumed that the characteristic of a valve is similar to a real pipe:

$$H_{Fi} - H_{Ti} = h_v = s_v q_v |q_v|^{r-1} \quad (8)$$

where h_v is the head loss of the valve, s_v is the resistance coefficient of the valve, q_v is the flow rate of the valve, and r is the index.

Therefore, the treatment of the valve is similar to that of a real pipeline from the hydraulic point of view. But, since there are many kinds of valves with different functional performances, there will be somewhat different program codes.

For example, a check valve is unidirectional; a limiter (similar to the above pump case) should be added in the calculation. For other valve styles, the corresponding limiters will be coded according to their characteristics in calculation programs.

4.3 Multi-source system

Multi-source herein refers to nodal points that have different constant pressure heads in the water networks. The system shown in Fig. 7 originates from the previous example in Fig. 1 but places an additional source point in the system. A commonly used technology is adopted to

deal with this case (Yan and Liu 2002). That is, an assumed or imagined water source (i.e., nodal point IP in Fig. 7) is introduced into the system, and all of the real sources are connected to this pseudo-source. Furthermore, this pseudo-source is the only source, rather than these real sources during the calculation process. Accordingly, some pseudo pipelines without resistance (no head loss) are used for these connections.

As a result, the multi-source system is transformed into a uni-source system again, and the aforementioned methodology in this paper can be used here. Also, the increased number of link pipelines will be the number of increased sources (excluding the original one). This can be described mathematically as

$$L' = M - N + S_0 \tag{9}$$

where L' is the number of link pipelines or fundamental loops for the transformed network, and S_0 is the number of sources.

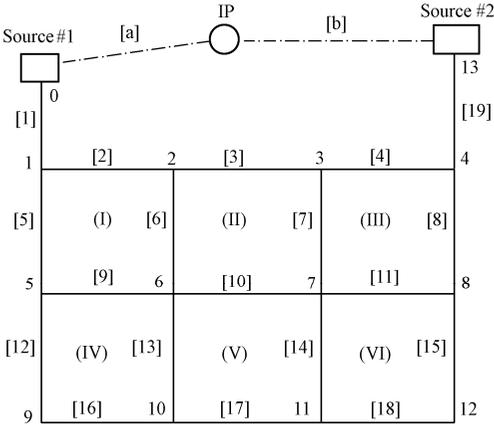


Fig. 7 Sketch of multi-source system

5 Practical applications

The study case was an urban water supply system of a highly modernized city in southern China. The simplified geometric construction is shown in Fig. 8. The daily water demand is about $8 \times 10^5 \text{ m}^3$. It contains 28 406 pipelines, 27 307 nodal points, and three water supply sources. There are three pump stations in the system to enhance the supply water head. The other essential data can be obtained from the central database with access by the program, and are omitted herein due to the limited space.

To construct a real-time central system for monitoring and scheduling the network, all the latest hydraulic and management information in this water system, including SCADA, GPS, GIS, LCD displays, and databases, need to be updated in a quasi real-time state with a short interval (e.g., 1 min) online. A fast and efficient hydraulic calculation algorithm is needed.



Fig. 8 Sketch of water distribution of southern city of China

In this system, according to Eq. (9), the total number of loops (or fundamental loops) is 1 102, which is much less than the number of nodal points (27 107). In this study, the H-C method, the N-E method, and the proposed S-T method were used to complete the hydraulic calculations for comparison. All three methods were completed on the same P-IV 2.0 personal computer. The results are shown in Table 1.

Table 1 Comparison of numerical results of different methods

Method	e_p (m)	Iteration times	CPU-time (s)	Accuracy (m)
H-C	0.001	187	82.5	0.001
N-E	0.001	106	56.5	0.000 85
S-T	0.001	51	17.1	0.000 23

Notes: e_p is the expected precision of pressure head and accuracy is the final obtained maximum error of pressure head.

6 Analysis and discussion

6.1 Result analysis

To make the comparison feasible and fair, the three results were calculated under the same systematic conditions. From the results in Table 1, the proposed spanning tree method is the best for convergence and efficiency, and with the other two it is difficult to satisfy requirements for real-time updating of information in this urban water system (both of them take longer than 1 minute).

The results of extensive tests with different expected precisions are shown in Figs. 9 and 10. The efficiency of the calculation is evaluated through CPU time consumption under certain precision conditions, while the accuracy is represented by the convergence speed of iteration. It is assumed that both the efficiency and accuracy have exponential distributions with the following forms:

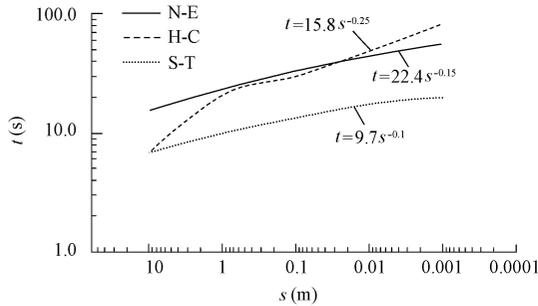
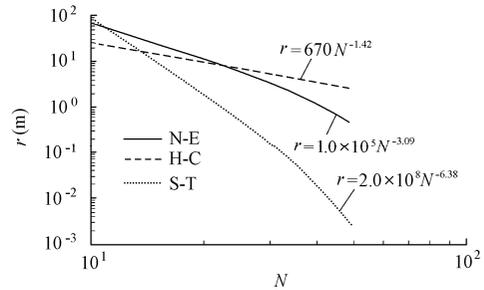
$$t = as^b \quad (10)$$

$$r = cN^d \quad (11)$$

where t is the CPU time consumption; s is the expected precision; r is the accuracy of calculation; N is the number of iterative steps; and a , b , c , and d are the coefficients.

Table 2 Calibrated coefficients

Method	a	b	c	d
H-C	15.8	-0.25	670	-1.42
N-E	22.4	-0.15	100 000	-3.09
S-T	9.7	-0.10	200 000 000	-6.38

**Fig. 9** Efficiency of three methods**Fig. 10** Accuracy of three methods

The fitted coefficients and plotted results are shown in Table 2 and Figs. 9 and 10. It is clear that the CPU time consumption for the proposed S-T method is much less than those of the two traditional methods. Its increasing tendency of time consumption with enhanced precision is also much slower than those of the other two. On the other hand, the N-E method has a lower efficiency than the H-C method during the initial stage, but it will exceed that of the H-C method gradually as the precision progresses further during later stages. This shows that the convergence of the traditional H-C method is less efficient in a high precision situation than that of both the S-T and N-E methods. Fig. 10 shows that the accuracy order of the S-T method reaches about 6.6, while those of the H-C and N-E methods are about 1.4 and 3.1, respectively, which demonstrates that the S-T method has a greater convergence speed with fewer iterative steps to obtain a certain precision.

The different performance in efficiency and accuracy of the three methods can be explained from the algorithm properties as follows:

First, the H-C method is a loop-equation based method. The number of loops in water networks is usually much less than that of nodal points or pipelines (Yan and Liu 2002). Therefore, the total number of equations to be solved is much less than that in the N-E method. This is the reason that the H-C method is more efficient in the initial stage than the N-E method.

However, the H-C method is based on the adjustment of the individual loop energy, and the adjusted error of each loop will propagate through the neighbor loops. As a result, several or many iterative steps later, when the calculated result in the network attains a relatively high level of convergence for the whole system, the error adjustment of each individual small loop has a great influence on the adjacent loops but has no effect on the big loops or the whole loop, resulting in many local convergences but nothing global (Yan and Liu 2002). Moreover, some adjustments of small individual loops even have negative effects on the total convergence

because of the linearized error of the loop equations, which always have an oscillatory effect on the final convergence.

The N-E method deals with the node energy equations through the nodal pressure head adjustment, and the pressure head just influences the flow rate of the pipes that directly connect to this node. Unlike the H-C method, this adjusted error of each adjustment will not propagate through the whole network, resulting in continual high efficiency during the whole calculation stage, especially during the later stages. Meanwhile, the sensitivity of the pressure head is much larger than that of the flow rate. That is,

$$\left| \frac{\partial H}{\partial q} \right| \approx ns|q|^{n-1} \ll 1 \quad (12)$$

Thus, during the later stages, the sensitivity of the hydraulic adjustment is more pronounced to the pressure head than that to the flow rate. Consequently, the N-E method will have a relatively higher and steadier convergence speed than the H-C method throughout calculation process.

On the other hand, the S-T method is based on the advantages of the H-C method. That is, it also solves the network through a loop-based equation to reduce the number of equations. The computation memory can be greatly reduced because it has fewer equations and variables than the N-E method. Furthermore, in the spanning tree, the fundamental loops all probably contain the small and large loops in the network (refer to previous sections). In actual fact, it decomposes the whole network into different-sized loops from the largest one to smallest one. During the later stages of the calculation, the effect of error adjustment is mainly from these large loops and can quickly lead to the final convergent results. As a result, the S-T method has a high efficiency and accuracy for practical networks.

Finally, through the generation of the spanning tree technique, the difficulties in initial estimations of the pipe flows in the traditional H-C method and the nodal pressure grades in the N-E method are reduced greatly in the S-T method.

6.2 Limitations

As seen through many trial tests for practical systems, the S-T method also has its own limitations in addition to the aforementioned advantages. One of most important is the generation of link pipes. If the link pipe is allocated by a non-main pipe or a very insensitive small pipe, it will influence the final efficiency and accuracy.

Fortunately, this can be resolved by our efforts in the preparation stage. Since the spanning tree is fixed during the whole calculation process after its initial generation (except for the updating of network information in the GIS during the calculations), the link of the tree can be generated by adding some essential limiters in the program, or, initially, software to assign the possible main pipe and/or sensitive pipe. Furthermore, the computational cost for this preparation is trivial, i.e., less than 3% of that for the main hydraulic calculation.

As a general efficient and robust methodology, the S-T method needs more practical examination and implementation in future work.

7 Conclusions

A new algorithm, the spanning tree method, has been proposed in this study for hydraulic calculations in large-scale practical water supply systems. The main principles and procedures of the formulation and computation are presented and implemented in this paper. Based on comparison with two traditional methods, the H-C and N-E methods, the proposed S-T method shows a relatively high accuracy and efficiency of hydraulic simulation for practical networks.

The properties of the S-T method related to its convergence and accuracy for large-scale system simulation are investigated through practical application. The important limitations and the practical implementations have been discussed in detail. More practical tests need to be conducted for this new method in the future.

References

- Bhave, P. R. 1981. Node flow analysis of water distribution systems. *Journal of Transportation Engineering*, 107(4), 457-467.
- Cross, H. 1936. *Analysis of Flow in Networks of Conduits or Conductors*. Urbana: University of Illinois.
- Deuerlein, J. W. 2008. Decomposition model of a general water supply network graph. *Journal of Hydraulic Engineering*, 134(6), 822-832.
- Duan, H. F., and Yu, G. P. 2006. Improved hybrid genetic algorithms for optimal scheduling model of urban water-supply system. *Journal of Tongji University (Science Edition)*, 34(3), 377-381.
- Gross, J. L., and Yellen, J. 2006. *Graph Theory and its Applications* (2nd Edition). Boca Raton: Chapman and Hall/CRC.
- Novak, L., and Gibbons, A. 1999. *Hybrid Graph Theory and Network Analysis*. Cambridge: Cambridge University Press.
- Ormsbee, L. E. 2006. The history of water distribution network analysis: The computer age. *Proceedings of the 8th Annual Water Distribution Systems Analysis Symposium*, 1-6. Cincinnati.
- Whaley, R. S., and Hume, R. 1986. An optimization algorithm for looped water networks. *Proceedings of the 18th PSIG Annual Meeting*, 1-19. New Orleans.
- Yan, X. S., and Liu, S. Q. 2002. *System of Water Supply and Drainage Distribution*. Beijing: China Architecture and Building Press.
- Young, B. 1994. Design of branched-water-supply network on uneven terrain. *Journal of Environmental Engineering*, 120(4), 974-980. [doi:10.1061/(ASCE)0733-9372(1994)120:4(974)]