Proceedings of the 10th Annual Water Distribution Systems Analysis Conference WDSA2008, Van Zyl, J.E., Ilemobade, A.A., Jacobs, H.E. (eds.), August 17-20, 2008, Kruger National Park, South Africa.

SPATIAL ANALYSIS TOOL FOR DEVELOPMENT OF LEAKAGE CONTROL ZONES FROM THE ANALOGY OF DISTRIBUTED COMPUTING

J. Sempewo[1,2], A. Pathirana[2,3] and K. Vairavamoorthy[2]

[1]Mityana District Local Government, P. O. Box 8761 Kampala Uganda

[2]UNESCO-IHE Institute for Water Education, P.O. Box 3015, 2601,
DA Delft, The Netherlands, Tel: +31(0)15 215 1854/Fax: +31(0)15 212 2921;
email a.pathirana@unesco-ihe.org

Abstract

Development of a methodology, rational and zoning scheme for the demarcation of complicated water distribution networks is important for making the zone demarcation process more accurate, economical, less time consuming, fast, repeatable, generic and optimal with respect to the cost of flow meters required. A new water distribution zone demarcation method is presented that uses the analogy of graph theoretic and graph partitioning principles used in distributed computing to distribute workloads among processors to suggest optimal zoning schemes based on balancing length, demand or flow within zones with the objective of monitoring of unaccounted for water. The method is developed into an automated prototype optimal zoning tool using C++ programming language, a graph partition tool used in distributed computing (METIS) and the EPANET tool kit. The tool is operated by a user interface written in the Python. Case studies are presented to demonstrate how the zoning tool is applied to the zone demarcation problem for the developed zoning schemes. The developed zone demarcation tool was observed to be an efficient and effective approach for the optimal demarcation of complicated water networks into optimal zones based on balancing length, demand or flow within zones. However the tool is sensitive to the number of partitions, the topology of the Water distribution network and the partitioning algorithms used. The tool can be used as a decision support tool for the optimal development and reduction of uncertainties in development of leakage control zones by decision makers. This will enable water companies to increase their productivities and also optimise resource allocations by reduction of the time to monitor, discover leaks and partition zones. This will lead to improved operating revenues

1. INTRODUCTION

Optimal demarcation of water distributions systems with the aim of monitoring unaccounted for water has become too complicated to continually thrive on trial and error, local knowledge, experience and “guestimation” which are in common use today. This is due to the fact that most of the existing designs were made without consideration that there would be district meter area zoning in future. The trend today is water distribution system upgrades especially in developing countries develop and enlarge gradually over time with little consideration to zoning criteria which makes the creation of District Meter Areas quite a challenging task. In addition water distribution networks are also characterised by suppressed low pressures due to pressure reduction campaigns and material replacement programs which also complicates the leakage location problem due to the reduced acoustic response from the leaks. These characteristics coupled together with the poor zoning methodologies have made the creation and management of zones and location of leakages to become hard tedious time consuming and costly. This is because the leak detection procedures have to be applied to almost each and every pipe in the network. This is further aggravated by the lack of a prototype optimal zoning tool.

In this paper a new water distribution zone demarcation method is presented that uses the distributed computing analogy in distribution of work loads among processors to suggest optimal zoning schemes for complicated networks with the objective of monitoring unaccounted for water. For example heavy scientific computations by parallel computers such as those that solve numerical solutions based on finite element methods require the distribution of the finite element mesh to the processors. This distribution is done with the objective of nearly equally distributing the work loads and at the same time minimizing the communication volume in order not to make the network the bottleneck .This is analogous to water distribution system zone demarcation problem where the different processors represent the water network zones. The water distribution network is first mapped onto a graph. Graph data is then discretised into a finite dimensional space through coding data of the graph properties to form concrete elements.

The objective of distribution of work loads equally among the processors in parallel computing is similar to that of creation of zones of a water distribution system which also has the objective of equally distributing loads in the zones. For the water network the loads are equivalent to the summation of any of the node or edge weights such as demand, flow or length obtained from network topology. One of the major shortcomings of applying classical graph theory to zone demarcation problem is the graph theories consideration of a network as a mere set of nodes connected by edges, without allocating for the `weights' in the networks. These weights can be nodal (e.g. Demands) and on edges (e.g. the diameters of inter-zone pipes or their flow). In the present model we explicitly allocate for the weights. In distributed computing, the objective of minimizing inter-processor communication (i.e. network traffic) aims to minimize the placement of adjacent elements to different processors. In the water distribution zone demarcation problem this is equivalent to minimizing the number of pipes that straddle two different zones. The paper proposes optimal zoning schemes for complicated water networks with the objective of monitoring un-accounted for water. These proposed zoning schemes together with the distributed computing phenomenon are developed into an automated prototype optimal zoning tool that can partition a complicated water network into optimal zones based on balancing length, demand or flow within zones. Performance of the tool is demonstrated by its application on a few complicated networks.

2. GRAPH THEOREM AND APPLICATION

Classical graph theory considers a graph as a mare set of nodes connected by edges without allocation for weights as illustrated in Figure 1(a).The major shortcoming with application of this theory to the zone demarcation problem is that it does not allocate for weights of the water network which is essential for the proper description and characterisation of the network graph as illustrated in Figure 1(b).These weights can be nodal (e.g. Demands) and on edges (e.g. the diameters of inter-zone pipes or their flow).

Classical graph (a) weighted graph (b)

Figure 1. illustration of weighted and un-weighted graphs

In order to apply graph theory to the zone demarcation problem there is need to extend the classical graph theory to allocate for weights of the nodes and edges. Most research regarding the application of graph theory in water distribution networks is in the area of graph theoretical modelling of the water network(Alonso et al., 2000; Gupta and Prasad, 2000).Deuerlein,2006 introduced a new decomposition concept of the common network graph according to its connectivity properties. However his concept did not allocate for weights of the graph. Further still little research effort has been made in the development of a water zone demarcation/sectorisation model tool that takes into consideration the weights of the graph. A new approach that explicitly allocates for weights in the solution of the zone demarcation problem is described. It is drawn from the methodology that has wide application in distributed computing for dividing work loads among parallel computations.

Analogy of Distributed Computing With the Zone Demarcation Problem

Parallel computations of large scale scientific computations such as the numerical solution of unstructured numerically heavy partial differential equations or finite element meshes is achieved through workload distribution amongst processors. This process of distribution of workloads among processors is what is called as distributed computing or parallel computing. To illustrate the distributed computing workload distribution phenomenon consider a problem domain in the form of a data structure composed of triangular finite elements that is to be solved by parallel computing. The problem domain is first converted into a graph problem with the nodes representing the unit of data and the computation to be performed. On the other hand the edges represent communication and dependence between the two vertices. The graph is discretized into a finite dimension space to form a concrete formula for a large but finite dimensional linear problem whose solution conversely solves the problem. The formulated concrete problem is then assigned to the computers for computation which distribute the workloads equally among the processors for task simplification. In addition the number of adjacent elements assigned to the different processors (communication volume) is minimised as an object.

Putting the water distribution zone demarcation problem in the analogy of the distributed computing phenomenon used in parallel computing, the data set to be simulated by the computers is equivalent to the water distribution network whose zones are to be ascertained. Conversion of the data set into a graphical structure is equivalent to decomposition and conversion of the water network into a graph before its demarcation. The different processors represent the different zones of the water network. The objective of distribution of work loads equally among the processors in parallel computing is similar to that of creation of zones of a water distribution system which also has the objective of equally distributing loads in the zones. For the water network the loads are equivalent to the summation of any of the node or edge weights such as demand, flow or length obtained from network topology. Minimisation of inter-processor communication aims to minimise the placement of adjacent elements to different processors. In the water distribution zone demarcation problem this is equivalent to minimising the number of pipes that straddle two different zones .This is similar to minimization of the diameter and pipes cut together with minimization of pipe flows that straddle two zones. It can be inferred that the efficiency of the outputs depend on how much the workloads are balanced and the loads are minimised which is also applicable for the zone demarcation problem.

To achieve the above objectives, parallel computers therefore strive to partition the computation (graph vertices) resulting into a graph partitioning problem where the objective is having equal partitions (zones) allocated to the processors and the constraint is minimization of communication volume (edge cut) Similarly since the zone demarcation problem is analogous to the distributed computing problem it can also be formulated into a graph partitioning problem where by graph partitioning algorithms applied in parallel computing can also applied to the zone demarcation problem.

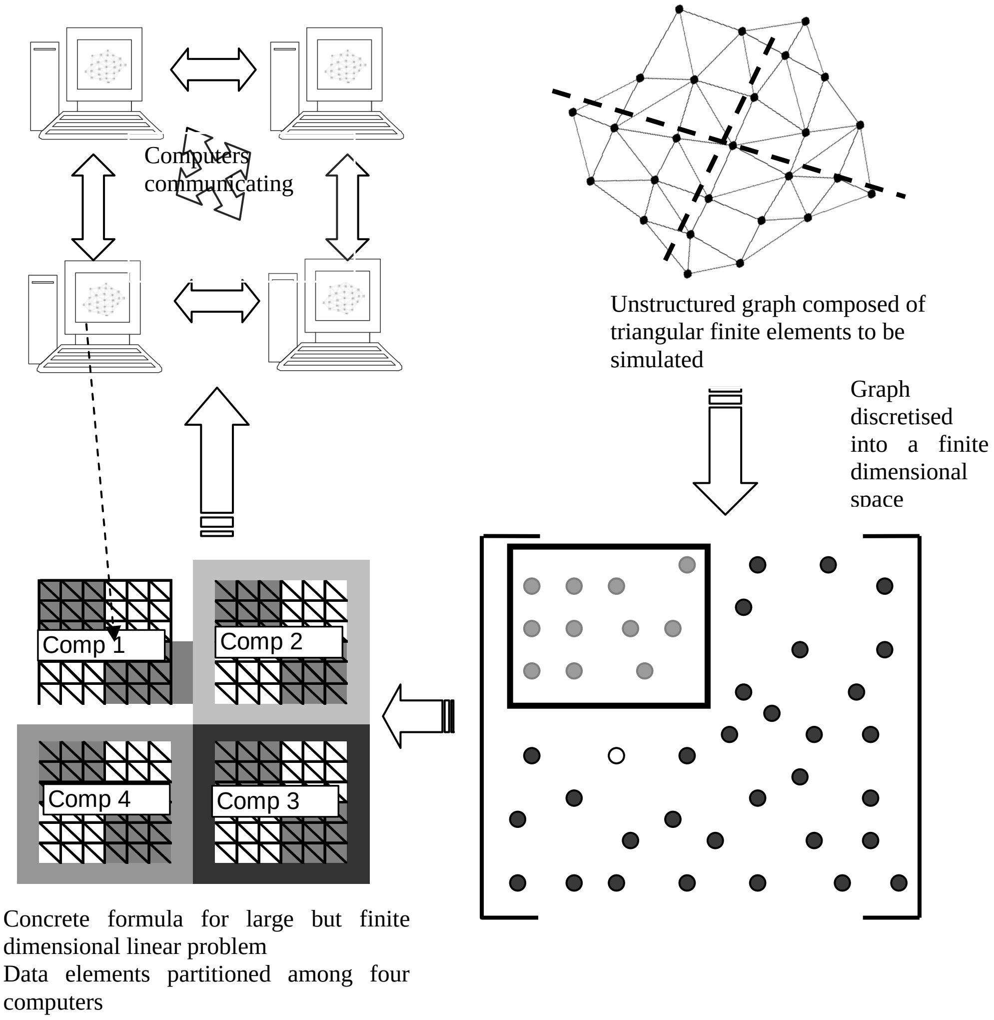


Figure 2. Distributed computing work load distribution phenomenon adopted for the zone demarcation problem

Similar to distributed computing load distribution problem where the first stage is the conversion of the problem domain into a graph problem, the water distribution network also requires the mapping and decomposition of the network into a graph to commence the solution of the zone demarcation problem. The decomposition concept of network graphs proposed by Deuerlein is adopted for the decomposition of the water distribution network into a graph where its vertices are nodes and edges are pipes or valves of the network (Deuerlein, 2006) .The graph mapped from the water network is thereafter discretized into a finite dimensional space through coding data of the graph properties to form concrete elements which are equally distributed among the processors. These elements through which the properties of the graph are stored are in the form of an array of nodes which are in a form a computer can easily process for partitioning. The process of the computer struggling to divide the computation evenly across a processor set while minimizing inter processor communication is therefore phrased as a graph partitioning problem.

3. THE ZONE DEMARCATION TOOL

A prototype optimal zoning tool has been developed that can partition a complicated water network into optimal zones based on balancing length, demand or flow within zones. Performance of the tool is demonstrated by its application on a few complicated networks. The tool was developed using C++, the METIS lib and the EPANET tool kit. Visual C++ was used to code the geometrical information of the EPANET input file format into a graph which was passed on as an input to the METIS lib partitioner. EPANET was used to compute the extended period simulations of the distribution system during the partitioning process from which weights were also generated.The structure of the developed model is as shown in Figure 3

The METIS Program and Library Interface

Publicly available software packages that execute the above partitioning algorithms like CHACO, METIS, PARMETIS, PARTY, SCOTCH, JOSTLE and S-SHARP are cited by (Fjallstrom, 1998; Schloegel *et al.*, 2000). METIS was selected because it provides high quality partitions, its extremely fast, and is able to partition graphs with weights. Further still it has partitioning routines that can be adopted for use in the partition of a graph (water distribution network) while balancing multiple constraints called Multi-constraint partitioning It is also able to minimize the total communication volume rather than minimizing the edge cut which improves the quality of outputted partitions. These algorithms are accessed through the stand alone METIS library and utilize unstructured techniques which are recommended for partitioning domains which hinge on geometry such as water distribution networks. (Karypis and Kumar, 1995b).

Zoning Schemes for Complicated Water Networks-Weighting Criteria

Zoning schemes were formulated from literature about DMA size determinants like served properties or a specified length (Yates and Donald, 2005) which for the purposes of this study are simplified to represent balancing demand and length respectively. For the edges the only weight assigned is diameter and for the vertices either balancing demand or length are assigned as weights for formation of the single objective zoning scheme. Parameters other than hydraulic parameters or parameters derived from the geometry of the distribution network were not considered in the development of the zoning scheme.Based on the above methodology and philosophy different zoning schemes are proposed and are as shown in Table 1

Table 1. Zoning schemes for complicated water networks

Zoning Scheme	Description of Prototype Scheme Class	Code -Argument in METIS programe	Edges (Pipes) Weight	Vertices (Nodes) Weight
1	No objective	0 0	-	-
2	Single objective	1 0	Pipe diameter as weight	-
3		0 2	-	Balancing demands in zones
4		0 1	-	Balancing Length in zone
5	Multi objective	1 1	Pipe diameter as weight	Balancing Length in zone
6		1 2	Pipe diameter as weight	Balancing demands in zones

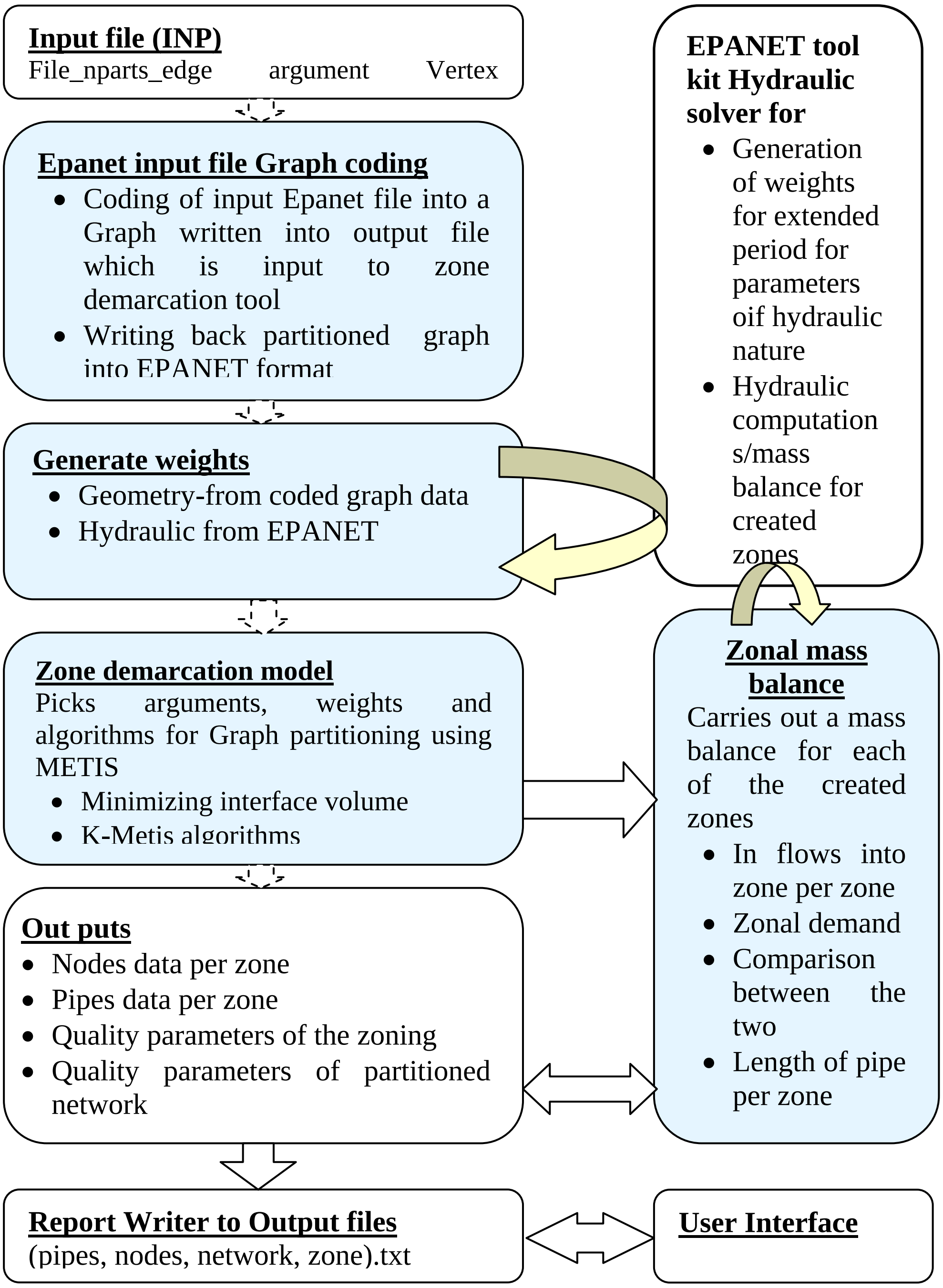


Figure 3 Developed Zone Demarcation Model Structure

The overall zone demarcation algorithm followed the structure of the model shown in *Figure 3*. It is as shown in the Figure 4.

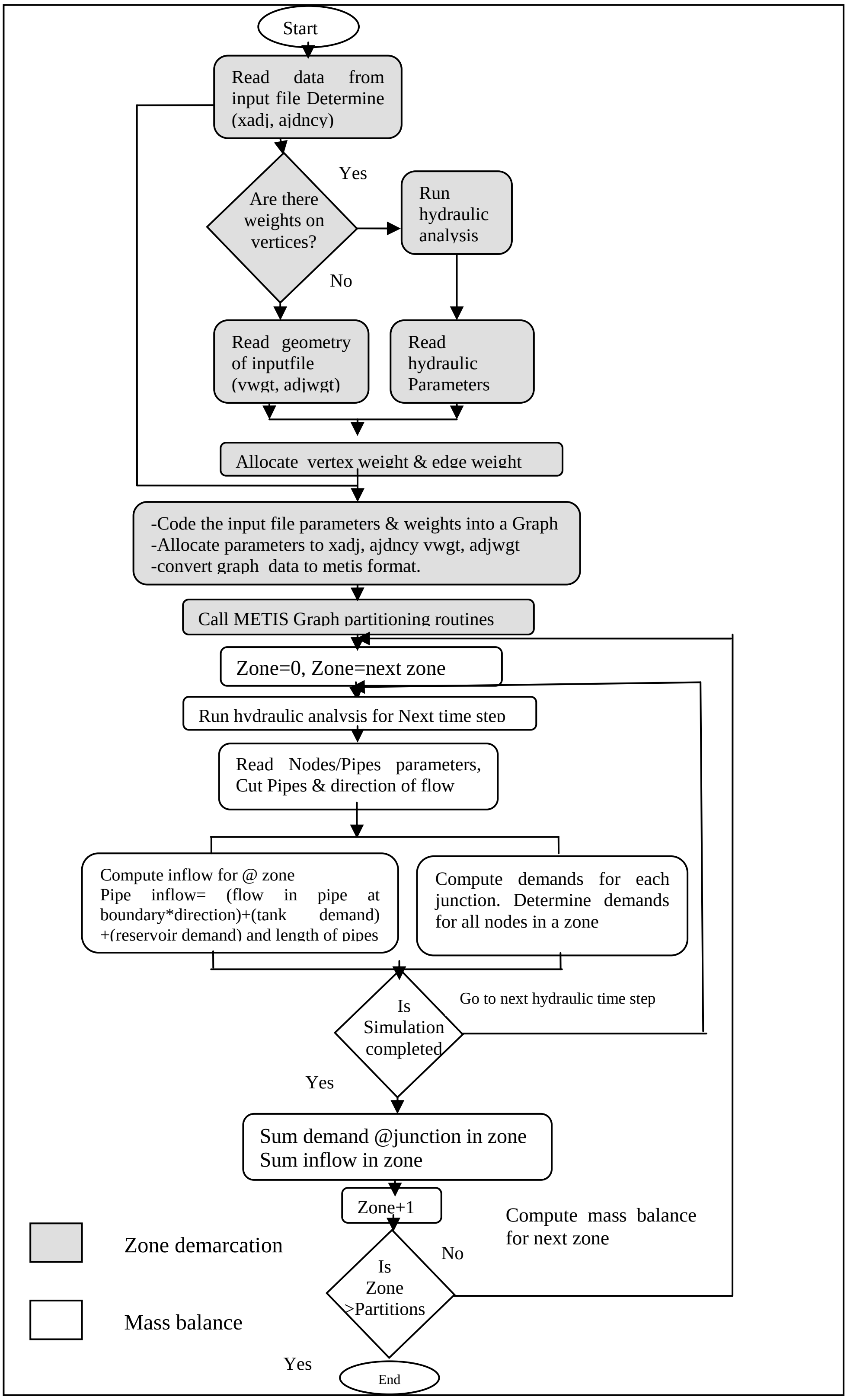


Figure 4. Zone demarcation and mass balance algorithm

Demonstration of Tool Performance on Test Network

For purposes of demonstrating how the tool works with respect to the developed zoning schemes, a small simple but tricky test network is developed that gives the partitioning tool tasks and complexity equivalent to those of large networks in performing the different partitioning objectives upon which the various testing scenarios are tested. For some test scenarios the test network is deliberately modified to create "predator-prey" zoning tool test scenarios developed to test and demonstrate the designed zoning schemes. The partitioning tool and proporsed partioning schemes are tested first on the test network results of which are shown and then later on real life cases. The developed test network consists of 12 pipes of varying diameters and lengths and 10 nodes. The test network has a total demand of 105l/s and a total length of 12000m. The schematic layout of the test network and the parameters are as shown in figure 5 and in Table 2.

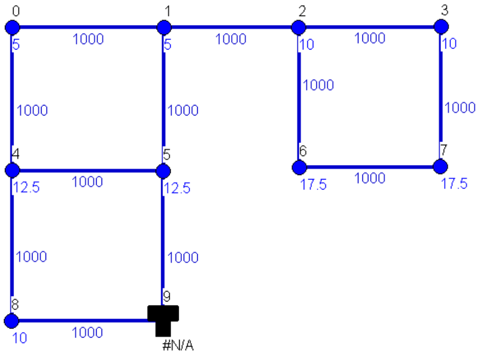


Figure 5. Test network

Table 2. Nodal and pipe data for test network

Pipe data

Pipe No	Start Node	End Node	Length	Diameter	C value
0	0	1	1000	500	100
1	1	2	1000	500	100
2	2	3	1000	500	100
3	0	4	1000	100	100
4	1	5	1000	100	100
5	2	6	1000	100	100
6	3	7	1000	100	100
7	4	5	1000	100	100
9	6	7	1000	100	100
10	4	8	1000	100	100
12	5	9	1000	100	100
13	8	9	1000	100	100

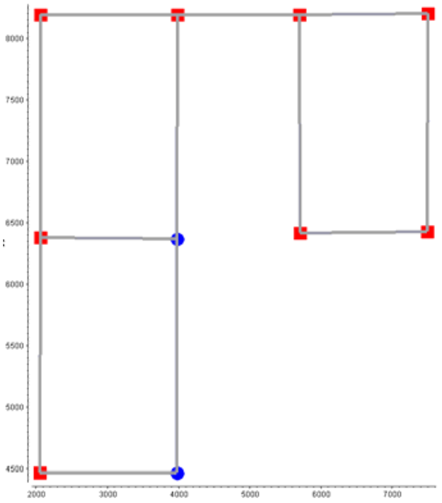
Node Data

Node No	Elevation	Demand
0	950	5
1	950	5
2	950	10
3	950	10
4	1200	12.5
5	1240	12.5
6	870	7.5
7	870	7.5
8	1450	35
9	2100	Reservior

A few case scenarios are used to demostrate the performance of the tool with respect to the different zoning schemes.

Scenario 1- Partitioning without due consideration of any constraint on the edges but while balancing lengths within the zones[0 1]

The test network is deliberately altered to allocate a length to one of the pipes equal to the total length of the rest of the pipes in the network (this is the Prey).On application of the partitioning toool on the network,the tool(predator) identifies the prey well as shown in placing all the other pipes in the network in one zone while placing the prey(deliberately altered pipe) in the other zone which is a good partition since the isolated pipe in the single zone is equivalent in length with all the other pipes in the network



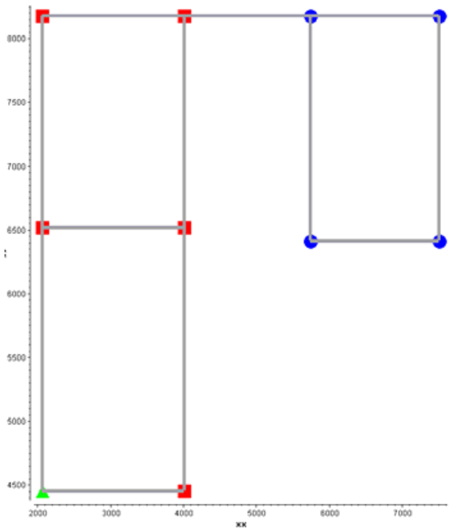
Legend/Symbol

○Circle 0 □Square 1

Figure 6. Testnetwork partition tool identifies deliberately altered test network pipe length in Scenario 1

Scenario 2-No constraint placed on edges while balancing base demands within the zones[0 2]

This scenario contrasts with scenario two in that the partitioning is done with the constraint of distributing basedemands eqaully among the zones and with no weights on the edges [2 0 2]. One of the nodes is also deliberately altered to have a base demand of 35l/s(which is 1/3 of the total nodal demand) and is placed (hiden)in two different nodal locations as prey.Partitioning the network into three zones produces three partions one of which has only one node with a base demand 35l/s . Figure 7 shows that the tool(predator) identifies the hidden node which is exactly 1/3 of the total base demand for each of the locations which signifies that the tool partitions the network according to its set objectives.



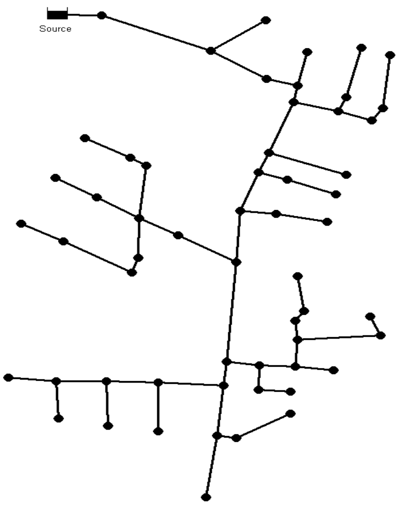
Legend/Symbol

○Circle 0 □ Square 1 △ Triangle 2

Figure 7. Scenario 2 test network partition tool out put results for prey identification of deliberately altered base demand

Zoning Tool Application

The zoning tool together with the graph partitioning algorithms were applied on several networks of varying topologies and complexity and proved carry out good partitions. The application of the zoning tool on the zone demarcation problem of water networks of different topologies is illustrated by considering a looped network with 29 nodes (SAFI) as shown in branched network input (GISTOWN) Figure 9 and a branched network with 58 nodes as shown in.For the evaluation of the performance of the tool with respect to node classification in a zone the percentage of nodes classified in the same zone was calculated and the results are shown in Table 3 .



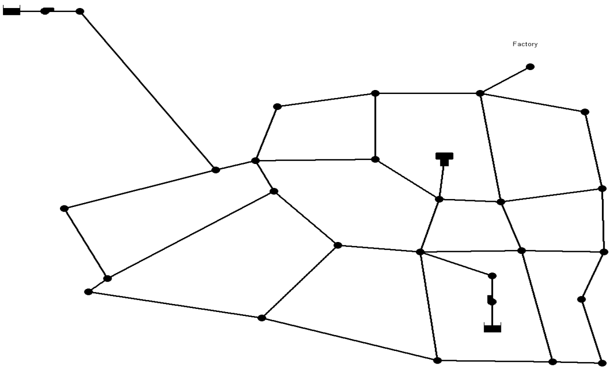
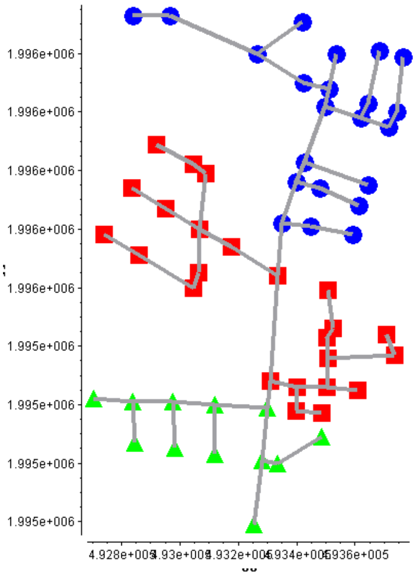


Figure 8. Branched network input (GISTOWN) Figure 9. Looped network input (SAFI)



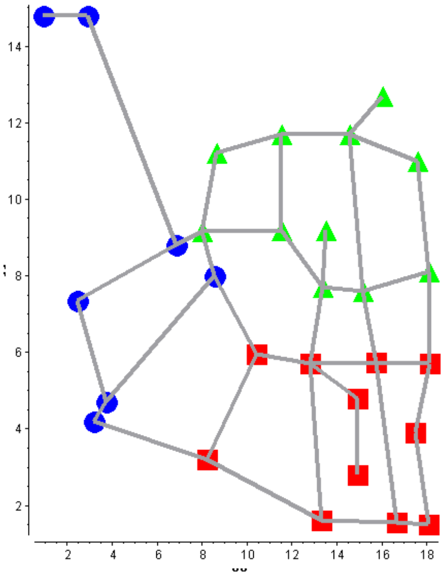


Figure 10. GISTOWN partitioned

Figure 11. SAFI partitioned

Performance of the Proposed Tool and Zoning Schemes

The sensitivity of the zoning tool to partitions, partitioning objectives and algorithms is illustrated by studying the tool outputs when applied on seleted networks of various topologies.For the evaluation of the tools sensitivity the percentage of nodes classified and the cut pipes Figure 12 are used as quality parameters to acces performance of the tool the results of which are shown in Table 3.

Table 3. Performance of zone demarcation tool with respect to node classification

NETWORK	DESCRIPTION	Min edgecut for Partitions less than 8						Min communication volume					
		No objective	Single objective		Multi Objective			No objective	Single objective		Multi Objective		
	Argument	0 0	0 1	0 2	1 0	1 1	1 2	0 0	0 1	0 2	1 0	1 1	1 2
GISTOWN Branched Network	Nodes classified in zone(%)	86	98	100	91	84	100	86	97	100	86	97	100
	Nodes miss classified in zone(%)	14	2	0	9	16	0	14	3	0	14	3	0
SAFI-Looped Network	% of nodes classified in zone(%)	100	100	100	100	100	100	97	100	100	100	100	100
	% of nodes miss classified in zone(%)	0	0	0	0	0	0	3	0	0	0	0	0

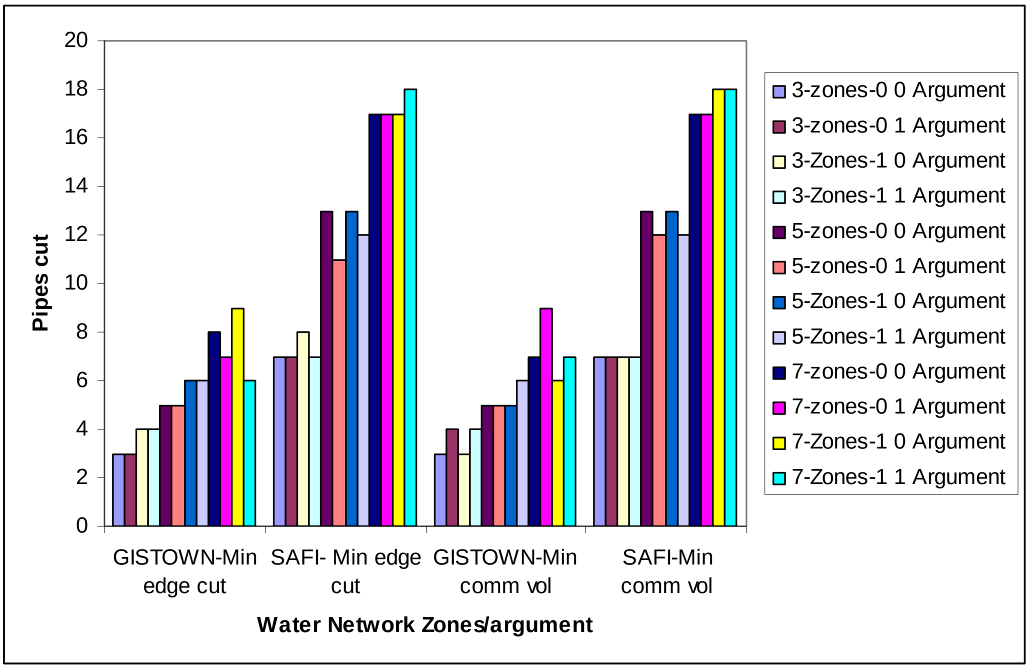


Figure 12. Pipes cut per zoning scheme per water network

It is observed from Figure 12 that the tool is sensitive to type of the network topology and complexity. Figure 12 shows that the quality outputs of cut pipes and constraint distribution are not sensitive to the type of partitioning algorithms as these tend to produce similar results for the same network. Partitioning of the networks becomes harder as the partitions increase as seen from Figure 12, partitioning quality decreases with increase in number of partitions which is analogous to the distributed computing case with respect to increase in partitions. Variation of the partitioning objectives within a zoning scheme affects the trends in the quality of the zones though not significantly. On the other hand variation from one partitioning schemes to another reduces the quality of partitions produced. It is also observed that the partitioning quality degrades as you shift from the no objective zoning scheme through the single objective zoning scheme to the multi-objective zoning tool.

Uncertainties and Limitations in the Tool

- The model tool gets uncertain especially for the multi objective zoning schemes as mentioned above and also for cases where the number of partitions is increased.

- The model does not take into consideration qualitative parameters of the network. Only quantitative parameters have been considered and these are the geometric and hydraulic parameters

3. LARGE CASE WATER DISTRIBUTION NETWORK APPLICATION

The developed zoning tool was also applied on a large complex network which is typical to a WDN today. The process was efficient and the tool produced good results as shown in Figure 13, Figure 14 and Figure 15

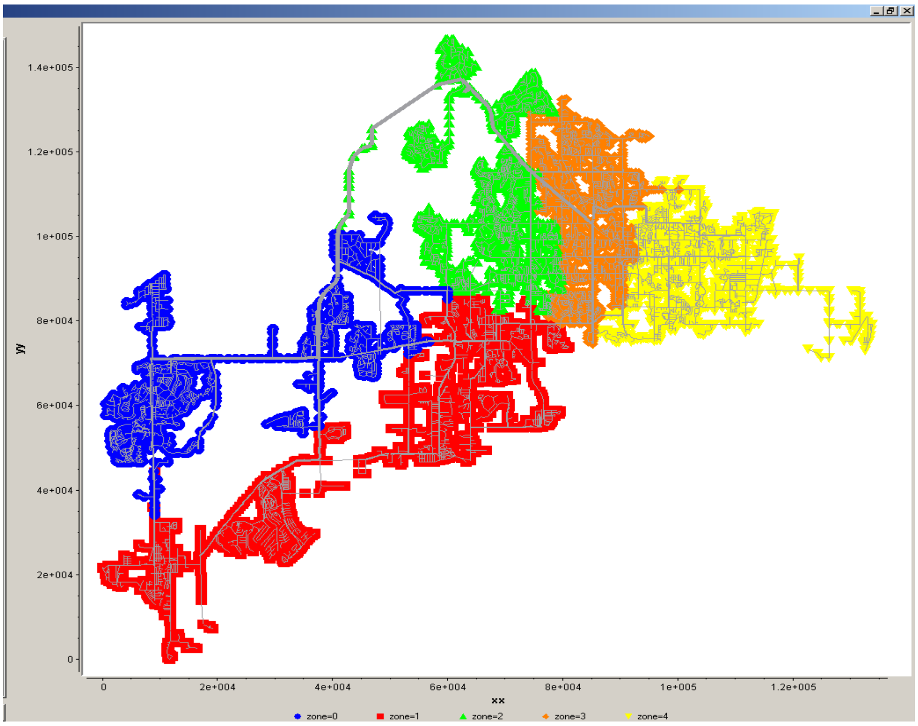


Figure 13. Zoning tool output on application on Large network (Argument 0 0) for five partitions with the objective of minimising edge cut using Hmetis

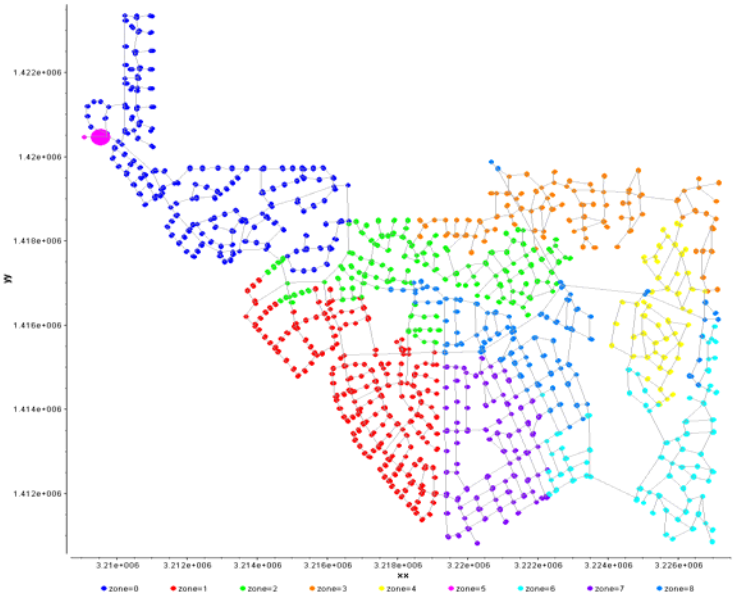


Figure 14. Zoning tool output on application on Wolf network (Argument 1 2) for nine partitions with the objective of minimising edge cut

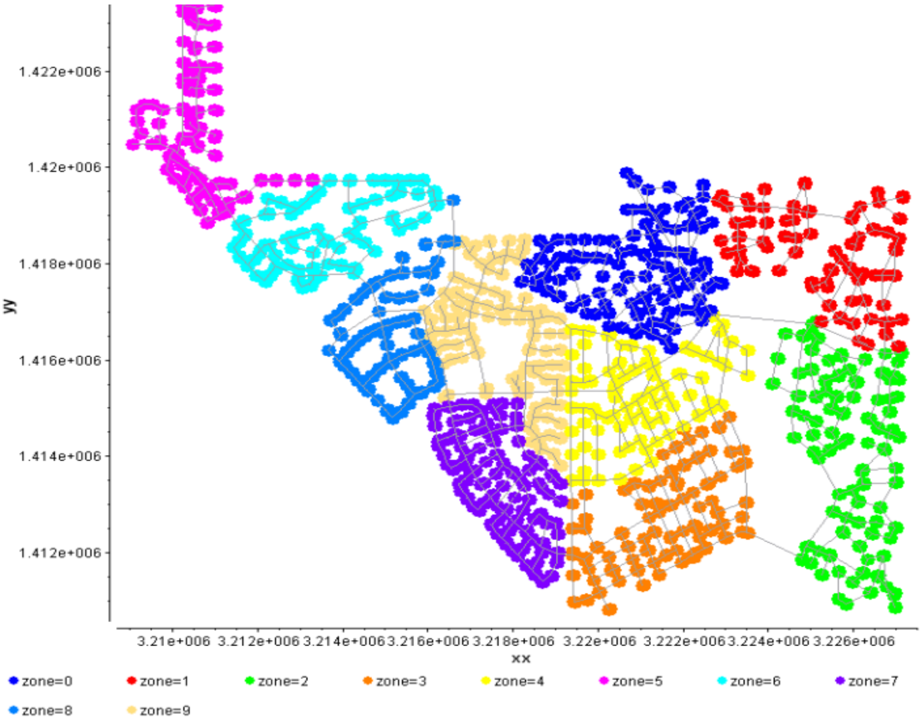


Figure 15. Zoning tool output on application on Wolf network (Argument 0 1) for Ten partitions with the objective of minimising communication volume

4. CONCLUSIONS AND RECOMMENDATIONS

In this study a spatial analysis prototype zoning tool; for the development of leakage control zones has been developed from the analogy of distributed computing using graph theoretic concepts. The tool has been tested on two networks of different topologies and complexities. The tool has also been applied on large case networks with thousands of nodes to verify it as a robust tool for use in the optimal zoning of large complicated networks. The developed zone demarcation tool was observed to be an efficient approach for the optimal demarcation of complicated water networks into optimal zones based on balancing length, demand or flow within zones. The tool mimics the distributed computing methodology of equally distributing workloads among processors and produces good contiguous partitions. Distribution of work loads equally among the processors in parallel computing is analogous to the zone demarcation problem in a water distribution system which also has the objective of equally distributing loads among the zones. For the water network the loads are equivalent to the summation of any of the node or edge weights such as demand, flow or length obtained from network topology.

The developed tool enables the consideration of weights which would have otherwise been not provided for in classical graph theory where a network normally is considered as a mere set of nodes connected by edges without weight considerations. The tool explicitly allocates for the weights on nodes (e.g. Demands) and on edges (e.g. the diameters of inter-zone pipes or their flow) in the water networks during the partitioning phase in form of zoning schemes. It should be noted that the developed tool only uses selected parameters as zoning schemes for creation of leakage control zones which are quantitative in nature although its agreeable that its not only quantitative parameters that should be considered when zoning of WDN. Further Research has to be done to holistically include all quantitative parameters and the extension of the model to cater for other zoning schemes that are qualitative in nature like pipe criticality, pipe condition with due consideration to the priority as these may vary from area to area.

From a business and industrial point of view especially for the enhancement of sustainability, the tool can help water companies increase their productivities and also optimise resource allocations by reduction of the time to monitor, discover leaks and partition zones. In addition the tool can be used to enhance leakage reduction and monitoring given that the network will have been broken down into smaller components that can be monitored

5. REFERENCES

Alonso, J. M., Alvarruiz, F., Guerrero, D., Hernandez, V., Ruiz, P. A., Vidal, A. M., Martinez, F., Vercher, J., and Ulanicki, B. (2000) Parallel Computing in Water Network Analysis and Leakage Minimization. *Journal of Water Resources Planning and Management*, **126**(4), 251-260.

Chamberlain, B. L. (1998) Graph Partitioning Algorithms for Distributing Workloads of Parallel Computations. *University of Washington Technical Report UW-CSE-98-10*, **3**.

Deuerlein, J. (2006) Efficient Supply Network management based on Linear Graph Theory. *8th Annual Water Distribution Symposium*, Cincinnati,Ohio,USA, August 27-30.

Fjallstrom, P. O. (1998) Algorithms for graph partitioning: A survey. *Linkoping Electronic Articles in Computer and Information Science*, **3**(10).

Gupta, R., and Prasad, T. D. (2000) Extented Use of Linear Graph theory for Analysis of Pipe Networks. *Journal of Hydraulic Engineering,ASCE*, **126**(No 1), 56-62.

Hendrickson, B., and Leland, R. (1995) The Chaco user's guide: Version 2.0. *Sandia National Laboratories, Albuquerque, NM*, 87185-1110.

Karypis, G., Aggarwal, R., Kumar, V., and Shekhar, S. (1999) Multilevel hypergraph partitioning: applications in VLSI domain. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, **7**(1), 69-79.

Karypis, G., and Kumar, V. (1995a) *Analysis of multilevel graph partitioning*, ACM Press New York, NY, USA.

Karypis, G., and Kumar, V. (1995b) MeTis: Unstructured Graph Partitioning and Sparse Matrix Ordering System, Version 4.0. *University of Minnesota, June*.

Karypis, G., and Kumar, V. (1998) *Multilevel algorithms for multi-constraint graph partitioning*, IEEE Computer Society Washington, DC, USA.

Karypis, G., Kumar, V., and (1995) A fast and high quality multilevel scheme for Partitioning irregular graphs. *Society for Industrial and Applied Mathematics*, **Vol. 20**(No. 1), 359-392.

Karypis, G., Kumar, V., Army High Performance Computing Research, C., and University of, M. (1996) *Multilevel Graph Partitioning Schemes*, Army High Performance Computing Research Center.

Schloegel, K., Karypis, G., and Kumar, V. (2000) Graph Partitioning for high performance scientific simulations. In: *Parallel Computing Hand book*, J. Dongarra, I. Foster, G.Fox, K. Kennedy, and A. White, eds., Morgan Kaufmann.