Drink. Water Eng. Sci. Discuss.,
doi:10.5194/dwes-2017-16-RC1, 2017

Drinking Water
Engineering
and
Science
Discussions
Open Access

Interactive comment on "Efficient Online Source Identification Algorithm for Integration within Contamination Event Management System" by Jochen Deuerlein et al.

Anonymous Referee #1
Received and published: 29 April 2017

General Comments:

The manuscript "Efficient Online Source Identification Algorithm for Integration within Contamination Event Management System" describes an algorithm for the identification of the location of a contamination event in a water distribution system. I think the paper is interesting and should be published, because (as highlighted by the Authors) existing software and approaches are not implemented for a real-time (or near-real time) control. Therefore, this manuscript provides a clear contribution to the current knowledge. I think the paper is in general well written and I don't have any major issue. However, I think that the paper could be clarified in some aspects (see below). While I understand that the Authors may have been limited (in terms of paper length) by the conference guidelines, I hope that the Journal will allow them to extend the manuscript, so that the content can be easily understood by a larger audience.

Specific Comments:

- Introduction, Line 27-28: I am not sure I understand this sentence. Do you mean that the algorithm is always running? Or it will be run (with a 1 minute time step) when there is an event?

The backtracking algorithm calculates the propagation of sensor signals in reverse time for both, negative (no alarm) as well as positive (alarm) signals. As result, the monitoring state of the observed network parts is always available. In other words for each location the time and the value of the last observation are regularly updated.

- Introduction, Line 28-29: I don't understand why there could be problems with the order of calculations. Could you please explain it? (Are the calculations performed on multiple processors or something similar? Or is it in case some of the inputs from the network are delayed?) Or maybe you can move this line in page 4 (about line 29) and give some details here.

The monitoring system in combination with hydraulic online simulation consists of several software components that share data using customized interfaces. In order to maintain proper workflow, the sequence of calculations follows specific rules. For example, the

backtracking algorithm requires the input of the current flow distribution and is therefore based on the hydraulic online simulation results. Deadlocks must be avoided.

- Page 2, line 25: please, explain or give an example of external time step. I think this refers to the time step of the steady state hydraulic simulation (in contrast with the time step used in the method of characteristics), but please, confirm.
Yes, it is true. The external time step is the one used by the hydraulic online simulation. However, it is not stead state simulation. The online simulation uses a slow transient model (rigid water columns model).

- Page 3, line 5: why are the computations performed even in case of no alarm? Maybe you can just refer the reader to section 4.2 (at least the reader will know that it will be explained somewhere).
The follwong text has been added:
“The latter is useful especially in th e case when flow directions change due to network operations. The continuous backtracking of negative sensor alarms allows the visualization of the current monitoring state of the system. For any location in the network the last time of observation is calculated. Due to the limitied coverage of the sensor network a contamination may not be detected by the sensor system because it is outside the covered area. It is important to be aware of this issue if other detections like customer complaints indicate an event contradicting the negative alarm states of the sensors.”
- Page 3, line 15-17: “In combination with the event driven method, which means that instead of time driven simulations only changes in water quality at the boundaries are considered, the memory requirements are minimized.” I don’t understand what this sentence means. Could you please reword it? In particular, I don’t understand the difference with time-driven simulation and event driven simulation
The paragraph has been rephrased:
“For implementation in SAFEWAER, an event driven method is used that strongly simplifies the common water quality equations: Instead of concentrations only binary quality states are calculated (contaminated/ non contmaninated) neglecting reaction and diffusion terms and using simplified mixing equations. An event is triggered every time a change in the (binary) boundary conditions occurs (for example start of intrusion in the forward case or release of a sensor alarm in the backward case). The algorithm sends a separator front through the system following the flow velocity of the water in the pipes (forward) or working against it (backtracking case). The moving front separates the network into regions that have distinct values for the binary quality state (contaminated / not contaminated). In the revers case a change in the sensor alarm state is considered as binary signal (node alarm -> alarm) Based on this simplifying assumptions memory requirements and calculation time are minimized in comparison with common water quality solvers. The fact that in case of real event the substance and input concentration of a contamination is presumably not known may serve as justification for the simplifying assumptions. “

- Page 3: line 25-26: “A specific weighting function has been developed that identifies the most probable locations for the contamination source based on the results of the backtracking method.” Could you please give a bit more detail and explain why you need a weighting function? Is it to take into account possible sensor failures/malfunctioning or the fact that the concentration may be very small and not detected in some pipes/sensors?

The following explanation has been added:
“A specific weighting function has been developed that selects a unique location for the contamination source out of the source candidates based on the results of the backtracking method and a worst case assumption. This second step is necessary since the backtracking algorithm in general does not give unique results, but the forward of calculation of the current spread of contaminant requires that a single source must be chosen from the source candidates. The region of possible source locations can be quite large depending on the efficiency of the sensor network.”

- Page 4 lines 21-22: “Positive alarms are generated by the EMS as soon as a pending alarm is acknowledged by the operator. After pressing the Acknowledge Alarm button [. . .]”. I think the software developed has the capability to automatically generate an alarm and start the computations. Maybe, at this stage, this is not performed in order to avoid several false alarms (or other reasons). However, I think it could be good if you could highlight that the operator response could be avoided.
The following explanation has been added:
“Alarms could trigger calculations automatically, but in order to avoid false alarms due to known events, at this state of the development every alarm has to be manually acknowledged.”

- Page 4, Line 24: could you please give more details about the choice of the worst case assumption? How is the single contamination source chosen? Is it the one that could affect the largest number of users (or the users with the largest demands or something else?). I think this is related to line 5 of page 5: does the backtracking algorithm always reach a water source? I think this depends on the location of the sensors and the contamination event: for example, if the contamination is after the source and you have a ‘negative’ sensor next to the source, the algorithm will select a pipe/node upstream of a ’positive’ sensor, but not the water source. I think you should clarify it in the manuscript. (I think you had this in page 8 lines 3-4, but I think it should be written (also) earlier in the paper).
The backtracking from all sensors reaches always a water source. The combination of positive and negative signals for all locations allows the correct identification of source candidates. The following sentence has been added:
“… states of all sensors including negative sensor alarms. Necessary conditions for source candidate locations are: 1) signals of all positive sensors passed the location during the backtracking, 2) no signal of a negative sensor was observed. Since the conditions are necessary but not sufficient the backtracking, in general, cannot give unique source locations.”

- Fig 1: maybe you could add some information about the test network (e.g. total pipe length), in order to give an idea of how big the network was. Also, you may be able to show in the figure the location of the closed valve (which is described in line 18 of page 5).
Done

- Page 8, line 10: Could you briefly discuss how these velocities are obtained? For example, are they computed from the hydraulic model based on some assumption on the demands? If so, this would introduce some uncertainty in the location of the contamination and the contamination zone. I understand that this is not the focus of this paper, but I think it should be mentioned. If the velocities are estimated/calculated in a different way, please, provide some more details (e.g. do you have a real-time

demand estimation algorithm?).
The following text has been added:
The continuous update of the model by online data aims at maintaining a sufficient match between the model parameters and the real situation in the filed for changing operational conditions. This is a step forward and improves the situation compared to calculations based on offline models. However, it must not be forgotten that uncertainty in model parameters still exists, especially the demands, affecting also the accuracy of the results of the source identification algorithm. The number of unknown parameters normally exceeds the number of measurements by magnitude.

Technical Corrections:
- Please, check the references: in page 2, they are referred as [7] and [8] instead of using the Authors' names. Also, I am not sure these references are reported at the end of the manuscript.
Done.

- Please, add Shang et al 2002 to the references.
Done.

- Eq 1a and 1b:
please, define PDE, IC and the symbols in the equation (ct, cx, and c0) - Please, define IVP (Initial Value Problem) in line 13 of page 2. - Page 8, line 16: I think the sentence should read "While the run time [. . .] is considered to be sufficiently short also for large networks, [. . .]"
Done.

Interactive comment on Drink. Water Eng. Sci. Discuss., doi:10.5194/dwes-2017-16, 2017.