Proceedings of the SWI 2014 Held in Delft

Monitoring the sewer system

Arnold Heemink[1],
Corine Meerman[2],
Sanjay Ramawadh[3],
Vivi Rottschäfer[2],
and Willem van Zuijlen[2*]

[1] TU Delft
[2] Leiden University
[3] Utrecht University

Abstract

The old city centre of Delft is sensitive to flooding caused by rainfall. We discuss the possibility and methods of combining data from observations, measurements and a theoretical model to determine the frequency and locations of these flooding events. We describe a protocol how the observations should be collected and present the determination of the measurement sensor positions as an optimisation problem. We also show some ideas about improving the presently used theoretical model.

KEYWORDS: Sewer system, modeling, monitoring, optimisation, knapsack problem, de st. Venant equations

ACKNOWLEDGEMENTS: The authors are grateful to Hans Korving (Witteveen en Bos) and Hai Xiang Lin (TU Delft) for discussions.

---

[*] Corresponding author: willem.van.zuijlen@math.leidenuniv.nl

1 Introduction

At the 98th Study Group Mathematics with Industry (SWI), held at Delft University of Technology from 27 – 31 January 2014 one of the questions was formulated by the company 'Witteveen en Bos'. Witteveen en Bos (from now on W+B) is a consultancy and engineering firm for water, infrastructure and environment. The problem they posed considers the sewer system of the city center of Delft. The old city centre is sensitive to flooding caused by rainfall. These events can cause sewage to flow from the sewer system back onto the streets and as a result the water on the streets will contain diluted (faecal) sewage [Man et al. (2014)]. Hence, the flooding does not only cause damage and leads to dangerous situations, with streets and buildings flooded, but also poses health risks due to exposure to contaminated water.

Inside the sewer, data is collected by using sensors that measure the water level. W+B already gathers data by using such sensors in the sewers in a district of the city Utrecht: Tuindorp. Their intention is to do the same in the city center of Delft. However, sensors are only able to measure the water level inside the sewer up to 30 cm below the surface level. When the water exceeds that level, there is no further information available and it is possible that water flows out of the sewer. In that case the sensors do not provide a accurate picture. Apart from obtaining data with sensors, W+B currently use a theoretical model to simulate the water levels inside the sewer. As in the case of the sensors, the theoretical model cannot predict the water level above surface. For these reasons, W+B desires to combine the known data below surface with data gathered above surface. W+B is interested in how this data above the surface should be collected and how it can be combined and compared with the data collected below the surface in the sewers. For this second question, our main objective is to detect at which locations and how often flooding occurs since this causes health threads. This will result in a complete picture of the sewer system. Another aim is to locate possible obstructions in the sewer system.

An obstruction in the sewer, such as a root intrusion or deposits, affects the sewer performance. These obstructions can be found using a camera survey. However, these surveys cannot be done regularly, since they are time consuming and expensive. The obstructions are not present in the theoretical model but they can be put in once their presence and position is known. This should be done on a regular basis since the occurrence of the obstructions changes over time. The data from the sensors gives more insight into the position of the obstructions and combining the sensor-data with data from above the surface will be very valuable.

Data that could be collected above the surface is for example photos of water (puddles) on the streets. At the moment, flooding incidents that are reported by citizens are already stored in a data-base. Increasing this number of reported incidents would improve the reliability of this data. In addition, we consider the possibility of making and collecting photos. One option that W+B is looking into is to use school children for this. Via these children, their parents and other citizens can also be motivated to gather data.

In Section 2, we present ideas and a strategy for collecting data above the surface. Considering where and when data has to be collected, we provide a protocol for the schools to follow in order to obtain sensible data. In Section 3 we describe a strategy for placing the sensors in such a way that the most important information is collected. Moreover, we optimise this placement by combining data from below and above the surface. In Section 4 we describe the theoretical model that is currently used to simulate water levels and give ideas for improving this model. In Section 5, our recommendations are summarized.

2 Strategies for obtaining data above surface

In collecting data above the surface, we have developed several methods which can be combined to get an accurate picture. Where possible, we would like the citizens to collect data. One route to increase the public awareness is via children and their schools. Previously, W+B had another program running at primary schools of Delft. Driven by this positive experience they are cooperating with primary schools to set up a program that is not only educative, but can also provide useful data about the sewer. We will discuss a way to implement such a program. Since primary schools have a regular schedule, this program might be missing the most important events, namely the heavy rain events. Therefore, we suggest to combine this with another program that can be used for just these events.

For both approaches we take into account a few aspects:

- Amount of time: How much time is available to gather and to evaluate the data?
- Place: At which places to gather data and on how many places?
- Information of a photo: What kind of information do we want a photo to contain?

After discussing these schemes we will give some suggestions of an app that aims to increase the amount of data collected voluntarily by citizens.

2.1 A program at primary schools.

Primary schools have a weekly timetable. For that reason, it is most likely that the sewer-program will included at a fixed time in the week. The idea is that the children spread over the city and take photos of puddles of water. However, it does not make sense to let them take photos if it has not rained lately. Therefore, we suggest to develop an alternative program as part of the sewer-program in case there is no water on the streets.

In case it has been raining lately, the photos should be taken at places randomly distributed over the city center. Since time and children are limited we do not expect to cover all of the city in one day. Therefore, there are multiple ways of choosing places, or, more specifically streets. Practical details, such as number and location of the schools, will point out whether it is possible to distribute the streets of interest uniformly or to have a circulating system over the schools such that different parts of the city at different moments are covered during the week.

Once the children have arrived at the (predetermined) streets they should take pictures of puddles of water on the street, and assign a measure of size to the puddle. In order to reduce the amount of data generated in this program – someone has to evaluate the pictures – a photo should only then be taken if there is water on the street. Even when puddles are not close to a manhole, the photos still contain useful information for the local government. To be able to link the photo to a particular manhole (or position), GPS information must also be included in the data. This program is summarised in the flow chart shown in Figure 1.

2.2 The approach in case of heavy rain events.

The method described above provides useful information on a rather regular base. However, the rain events that are very important, those of heavy rain or storms, might not be covered. These storm events could give the most valuable information, since flooding is most likely to occur during such an event. Therefore, the above school program should be complemented by another method for collecting information from these events. For this, one or more persons should be readily available.

Weather predictions can be used to determine the moments at which data should be collected. Data has to be gathered from the moment it starts raining until the moment that the water has disappeared from the streets.

We select streets where data should be collected. We select the streets with the use of the sensor data, for example, by choosing streets where the sensors indicated that the water level gets to a high level. Furthermore the places where an incident was reported will be visited. Then if there is a puddle, the person will take a photo and/or assigns a measure of size to the puddle and repeatedly come back until the water has disappeared.

Similar to the previous approach, GPS coordinates must be included. Moreover, in case an estimate of the amount of water on the street can be

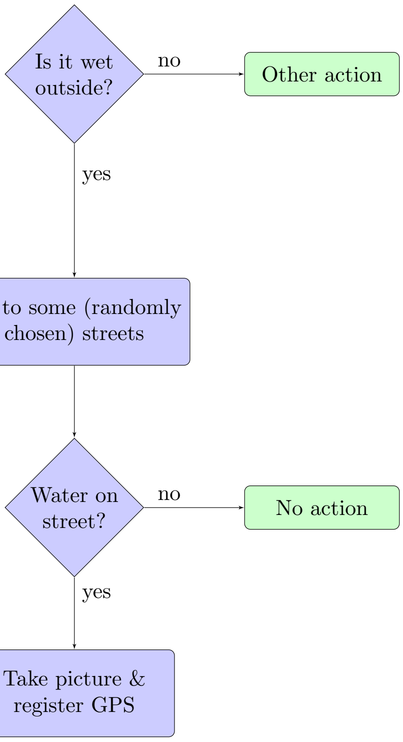

Figure 1: Flowchart of the program for school children.

given, this information should be added. The outline for this program is summarised in the flow-chart shown in Figure 2.

2.3 Suggestions for an app/website

At the moment, people can contact the local government by phone for incidents. We suggest that, apart from this, people should be able to report incidents via other media like internet. Incidents reported by phone-calls are already useful in order to detect possible defects or obstructions in the sewer system. However, information of phone-calls is rather subjective. More objective information can be obtained via a website or app that enables the persons to add extra information to the street name where the incident is mentioned, like pictures with GPS data and other observations. In such a way, one can also provide a guideline of how the person should include the data. By presenting such a guideline and an easy way to report incidents (without having to make a phone-call), we expect the number of reports to increase and to

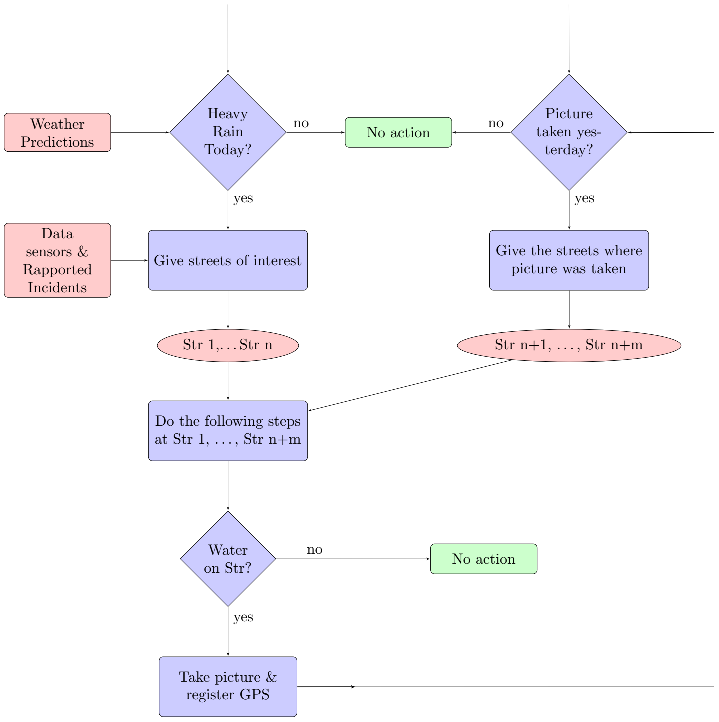

Figure 2: Flowchart of the program for heavy rain events.

be of better use. For this approach, the citizens should be made aware that their reports are very valuable and that they can help to reduce flooding in the future.

3 Placing of the sensors

In order to monitor the behavior of the sewer system below the surface efficiently, we want to place sensors at those manholes which are most likely to overflow. These sensors measure the level of the water in the sewer once every minute. Moreover, every sensor has a threshold up to 30 below the surface level. This threshold is known by W+B. The sensor data is not only useful for showing when and where a manhole has flooded, it can also be used to find obstructions in the sewer system. The method for finding these obstructions based on sensor data is also known by W+B [Bijnen et al. (2012)].

Next to the measurements with sensors, data is also available from reported incidents and photographs by citizens. We want to combine these data with the sensor data to determine those manholes which are most likely to be flooded; the problem spots.

A maximization problem is set up to find the optimal placement of the sensors. The sewer network can be seen as a graph $G = (V, E)$, where the set of vertices V corresponds to the manholes and the set of the edges E corresponds to the pipelines. Let us define $n := |V|$. The main assumption will be that the sewer network contains no obstructions. We start with a clean network which, besides containing no obstructions, also has no sensors placed yet. W+B has a theoretical model of the sewer network, which can accurately simulate the level of the water below the surface in a clean network. Using this theoretical model, a well-educated guess is made for the initial placement of the sensors. Several simulations are done with the model, using different values of rainfall as an input. A sensor will be placed at those vertices where flooding is most likely to happen. If the well-educated guess requires placing less sensors than we have, we can place the remaining sensors at random locations.

3.1 Setting up the problem

The idea is to determine a placing of the sensors such that only the measurements of the sensors can give us an accurate view about the places of the incidents. We expect that the placement at the start is not suitable enough for this, so we want to replace certain sensors. We assume that the sensors are working properly, since malfunctioning sensors can be detected through data validation. If a placed sensor indicates that the threshold has been reached and at the same time there is observed data about an incident, then the sensor should not be removed. If the observed data shows no incident, then it does not matter whether the sensor is removed or not. If a vertex does not have a sensor and observed data shows an incident, a sensor should be placed there. If the observed data does not show any incidents, it does not matter whether the sensor is placed or not.

Each vertex will be given a value, which depends on whether or not the corresponding manhole has a sensor and also on the data available about this manhole. The value of a vertex indicates how likely it is that the corresponding manhole overflows; the higher the value, the more likely it is to happen. The value of a vertex will be denoted by $\alpha_{t,l}$, where t stands for the day of the measurement and l represents the corresponding manhole. The value is determined as follows. First, assume that the corresponding manhole l has a sensor. Then, for each day t, the value $\alpha_{t,l}$ equals 1 if the sensor reaches the threshold on day t as well as there is an incident reported on day t for this manhole. Otherwise, the value is 0. Now, assume that the corresponding manhole l has no sensor. Then, for each day t, the value $\alpha_{t,l}$ equals 1 if either there is an incident reported on day t for this manhole, or if there is an incident reported for some other manhole nearby in the network and at the same time a simulation of the rainfall in the theoretical model indicates that this manhole has flooded as well.

We assign, to each vertex l, a variable x_l which can only attain the values 0 and 1. The expression $x_l = 1$ means that we want to place a sensor at the manhole corresponding to vertex l. The expression $x_l = 0$ means that we do not want to place a sensor there. The variables x_l are subject to some constraints which arise from practical considerations. For example, sensors should be maintained regularly, which comes with a cost. Moreover, these maintenance costs might be different depending on the location of the sensor. To include this constraint, we define, for each vertex l, the variable w_l which is the average maintenance cost per day (or any other timespan) for a sensor at the manhole corresponding to vertex l. The sum $\sum_l w_l x_l$ then equals the average maintenance cost per day to place or remove sensors, which is likely required to be smaller than some predetermined constant W, the daily budget. Another constraint is given by the limited number of sensors. Since $\sum_l x_l$ equals the number of sensors we want to place, we require this sum to be smaller than some other predetermined constant C, the maximal amount of sensors we can place.

In order to determine the problem spots, that is, the vertices corresponding to those manholes which are most likely to overflow, we use the data gathered over a certain time frame, for example a year. For each vertex l, the sum $\sum_t \alpha_{t,l}$ is the total value of the vertex l in this timeframe. The quantity $\sum_{l\in V}\sum_t \alpha_{t,l}x_l$ now represents the total value of the vertices where a sensor should be placed. Since we want to place the sensors as efficient as possible, this total value needs to be as high as possible. The maximisation problem can now be written as:

$$\max\left\{ \sum_{l\in V}\sum_t \alpha_{t,l}x_l \;\middle|\; \begin{array}{l} x_l \in \{0,1\},\ l = 1,\cdots,n \\ \sum_{l\in V} w_l x_l \leq W \\ \sum_{l\in V} x_l \leq C \end{array} \right\}. \tag{1}$$

3.2 Solving the problem

The optimisation problem stated in expression (1) is equivalent to the 0/1 knapsack problem, which is a known problem in combinatorial optimisation. The 0/1 knapsack problem with one constraint is formulated as follows. Suppose we have n objects, each with a value v_l and a weight w_l. The objective is to choose objects in such a way that the total weight of the chosen objects does not exceed some predetermined constant W and, moreover, that there is no other selection of objects satisfying the same weight constraint which has a higher total value. The difference between the general knapsack problem and the 0/1 knapsack problem is that, in contrast to the general knapsack problem, each object can be chosen either once or not at all. The 0/1 knapsack problem with one constraint is mathematically formulated as follows:

$$\max \left\{ \sum_{l=1}^{n} v_l x_l \;\middle|\; \begin{array}{c} x_l \in \{0,1\},\ l = 1, \cdots, n \\ \sum_{l=1}^{n} w_l x_l \leq W \end{array} \right\}. \tag{2}$$

The general 0/1 knapsack problem, that is, the 0/1 knapsack problem with any number of constraints, is an optimisation problem that is known to be NP-complete [Garey and Johnson]. All NP-complete problems are characterised by the fact that there is no known algorithm that solves these problems quickly. However, there are several heuristic methods known which find a solution close enough to the optimal solution. In the case of the general 0/1 knapsack problem, the most commonly used method is dynamic programming [Salkin and Kluyver (1975)]. In the specific case where the 0/1 knapsack problem is of the form in equation (2), with all weights equal and, without loss of generality, all equal to 1, the problem belongs to class P and can be solved exactly. The solution is $x_l = 1$ for the $\lfloor W \rfloor$ objects with the highest values, with $\lfloor W \rfloor$ the largest integer smaller than W. So the maximisation problem in (1) coincides with this specific case if the average maintenance costs per day for all sensors are equal, as the two constraints can be reduced to one.

4 Mathematical modelling of sewer networks

In this section we give a brief summary of a model that describes the dynamic behaviour of the water level in a sewer system. For a detailed description the reader is referred to [Cunge et al.]. We also discuss the possibility to use the data that will be collected above the surface and by the sensors in the sewer system to improve this model. Furthermore we will suggest a method to identify locations where obstructions might be located.

A model for open channel flow (developed originally for river systems) is based on the so-called "de st. Venant equations", also referred to as the "1D shallow water equations". This model can also be used for sewer systems. The equations are respectively derived from conservation of momentum and conservation of mass:

$$\begin{aligned} &\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + g\frac{\partial h}{\partial x} + gS_l - \mu S_f = 0, \\ &\frac{\partial h}{\partial t} + \frac{1}{b}\frac{\partial (uA)}{\partial x} = S, \end{aligned}$$

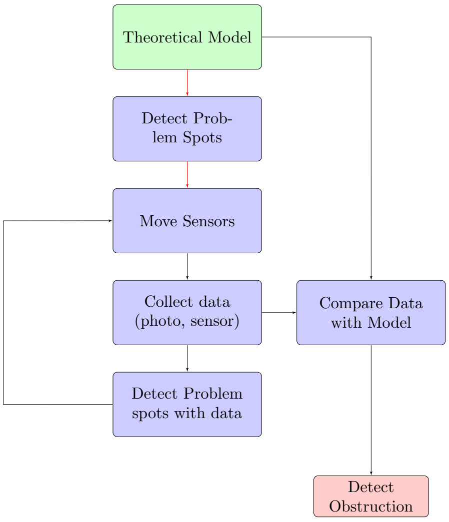

Figure 3: Flowchart summarising the recommendations about moving the sensors to a more optimal position and also about finding possible obstructions in the sewer network.

where x is the spatial coordinate (in one dimension along the sewer pipe), t is the time and

$$\begin{aligned}
g &= \text{acceleration of gravity},\\
u(x,t) &= \text{water velocity},\\
h(x,t) &= \text{water level},\\
S_l(x) &= \text{slope of the channel},\\
S_f(x,t) &= \text{friction term},\\
\mu(x) &= \text{friction coefficient},\\
b(x,h) &= \text{width of the channel},\\
A(x,h) &= \text{cross sectional area},\\
S(x,t) &= \text{water source or sink}.
\end{aligned}$$

The sewer system consists of pipes that are connected at junctions. The velocity and level of the water in each of these pipes are described by the above model. At the network junctions the amount of water that flows in has to be equal to the amount of water that flows out. In other words the total inflow is equal to the total outflow at a junction. This leads to one boundary condition for the partial differential equations described above. Since we still need one other boundary condition at the junction we could in addition assume that the water level is continuous at the junction. This boundary condition does however not take into account the complicated flow pattern that may occur at a junction in case of a sharp curve. Therefore one often assumes continuity of the quantity

$$h + \alpha \frac{u^2}{2g}.$$

Here the second term is introduced to account for local effects at the junction due to the bulk motion of the fluid. This effect creates small differences of the water levels in the various pipes at the junction: The higher the velocity at the entrance of the pipe, the lower the water level at this location will be. The empirical factor $0 \leq \alpha \leq 1$ is a tuning parameter. In order to obtain a numerical model for the sewer network, the well-known Preissmann scheme [Cunge et al.] can be used. This scheme is very attractive for networks since it can deal easily with non-equidistant grids and is also capable to include the boundary conditions just described.

In the model S is the water source or sink and it represents the amount of water that enters or leaves the sewer network at the pipe (when S is negative, water flows out of the sewer network). If S is 0, then the total amount of water in the system is preserved. One of the difficulties in using this model is that there is hardly any information available about the water source or sink S. In the sewer system the water level can exceed the level of the street and then water will flow out of the system. But generally the amount of water that flows out of the sewer system is not known. Hence a good estimate of S is desirable. We suggest to explore the possibility to use the information that is gathered on street level (like pictures) together with the model simulations to estimate the parameter $S(x,t)$ in various sections of the sewer network model. This calibration procedure can be formulated as an optimisation problem by defining a cost function that measures for a given parameter the difference between the model results and the available measurements. Here usually the least squares criterion is taken as cost function. One value of this cost function for a given parameter (for S) can be computed by running the model with this parameter value and by comparing the results of the model at the sensor locations with the measurements. Using an optimisation scheme the parameter can be improved step by step and finally the best estimate of the parameter is the parameter for which the cost function is smallest and therefore for which the model simulation is as close as possible to the measurements available.

As we indicated before, water level measurements with sensors combined with information from photos can also indicate possible locations for obstructions. The available data can be combined with the model simulations in order to estimate the friction coefficient $\mu(x)$ at various locations in the network by means of the calibration procedure just described. If the estimated value for the friction coefficient is significantly larger than the original value in the model at a certain location in the network, this is an indication that there might be an obstruction.

5 Recommendations

Here, we give a summary of the recommendations regarding collecting data, moving sensors and adjusting the theoretical model.

To collect data above street level we suggest to implement a program for the primary schools in Delft to collect photos of puddles of water in the streets (see also Figure 1). Apart from that, we suggest to use a complementary program to collect data in case of heavy rain events, described by the flowchart in Figure 2. Additionally to the two programs, we suggest to develop an easy app and/or website on which citizen can report their observations of a flooding incident.

We suggest to use the theoretical model as the basis for the initial placement of the sensors. This is done by simulating many different rainfall events and by choosing problem spots which are most likely to overflow. Once the sensors are placed, data can be collected from above the surface as well as from below the surface. After a certain, not necessarily predetermined, time frame, we use all the gathered data to optimize the positions of the sensors according to the reduced knapsack problem. This procedure can be repeated at any time. The flowchart in Figure 3 summarizes our recommendations.

For the theoretical model we suggest that the amount of escaping water could be estimated using the data from above the surface. Furthermore via an estimate of the friction coefficient that fits the measured data, one might be able to locate the obstructions by comparing this estimated friction coefficient to the normal friction coefficient.

References

M. van Bijnen, H. Korving, and F. Clemens. Impact of sewer condition on urban flooding: a comparison between simulated and measured system behaviour. *Proceedings of the Ninth International Conference on Urban Drainage Modelling*, 2012.

J. A. Cunge, F.M. Holly, and A. Verwey. *Practical aspects of computational river hydraulics.* Pitman Advanced Pub., Boston.

M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* Freeman and co., New York.

H. de Man, H.H.J.L. van den Berg, E.J.T.M. Leenen, J.F. Schijven, F.M. Schets, J.C. van der Vliet, Knapen F. van, and Roda Husman A.M. de. Quantitative assessment of infection risk from exposure to waterborne pathogens in urban floodwater. *Water Research*, 48:90–99, 2014.

H.M. Salkin and C.A. de Kluyver. The knapsack problem: A survey. *Naval Research Logistics*, 22:127–144, 1975.

Proceedings of the SWI 2014 Held in Delft

, , , , ,