Applications of Machine Intelligence Methods for Near Real-Time Flood Modelling

Dragan A. Savić, Albert S. Chen, Michele Guidolin, Rebecca Austin, Michael J. Gibson, Edward C. Keedwell, Slobodan Djordjević, Andrew P. Duncan, Bidur Ghimire
Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences, University of Exeter, United Kingdom

Urban flooding has been estimated to cause £270 million damage annually in England and Wales, currently affecting some 80,000 homes, rising up to 300,000 - 400,000 per year by the 2080s (OST, 2004). The frequency and impact of urban flooding are expected to increase due to deteriorating infrastructure, and new development, an increase in paved areas and climate change. Although institutional arrangements have been identified as a major obstacle in dealing with urban flooding in a recent study by Defra (Gill, 2008), the urgent need for improved modelling tools has also been highlighted. Modern mapping technologies, such as synthetic aperture radar, aerial digital photogrammetry, and light direction and ranging (LiDAR), have significantly improved data availability and accuracy for flood modelling applications. However, existing simulation tools are unable to deal efficiently and effectively with flood modelling over the large urban areas using high-resolution terrain model, which are critical to describe micro features that affect flood propagation in the urban environment. In particular, computational efficiency is an issue when models have to be used in flood risk/uncertainty/mitigation analyses, for which repeated simulation runs are required. Furthermore, these models are unable to provide results in real or near-real-time, which is essential for effective flood warning systems in urban areas.

The Cellular Automata Dual-DrainagE Simulation (CADDIES) project, funded by the UK Engineering and Physical Sciences Research Council, aimed to produce fast and accurate algorithms for flood modelling in urban areas with reduced-complexity models and modern computing acceleration techniques. The objective is to improve the speed and efficiency of flood modelling that involves both urban surface and drainage system (2D urban surface flow and 1D sewer flow) for real/near-real time applications.

Cellular automata (CA) are dynamical systems that were originally developed as a conceptual model of the Universe and its laws and have been shown to simulate highly complex physical systems (von Neumann, 1966). More recently, they have been used for describing models of systems in which many simple components act together to produce complicated patterns of behaviour (Yeh and Li, 2007). CA models are normally discrete both in space and time, with space represented by a grid made up of many cells, whose states may change between time intervals. States are updated in accordance with fixed rules that depend on a cell’s value and the values of its neighbouring cells. Due to its discrete nature, any physical system satisfying differential equations may be approximated as a CA, by introducing finite differences and discrete variables. Cellular Automata offer a versatile method for deriving reduced complexity models of physical systems. Since the computations are limited to neighbourhoods and the local rules are identical for the whole lattice, CA’s have proved to have an inherent massively parallel computation capability. This advantage can be readily exploited using the latest General Purpose Graphics Processor Unit (GP-GPU) hardware installed in standard PC’s together with software tools such as Compute Unified Device Architecture (CUDA) (nVIDIA, 2015).

Genetic programming (GP) is another evolutionary computing method, which can generate computer programs that perform a user-defined task (Gibson et al., 2014; Koza, 1992). GP builds on methods derived from genetic algorithm (GA), but whereas GA are generally used to evolve the best values for a given set of model parameters, GP allows the basic structure of the model to evolve, together with the values of its parameters. This feature of GP is quite promising for evolving symbolic CA rules. Giustolisi and Savic (2006) developed an advanced GP methodology, Evolutionary Polynomial Regression (EPR), which combines numerical and symbolic regression. We will extend the EPR methodology to allow appropriate choice of primitive mathematical functions to be used as building blocks for evolving CA rules. This can be done in such a way that evolved CA rules could be anything from purely black-box models to fully white-box models.
Typical flood simulations use full hydrodynamic models, which are computationally expensive at finer resolution, and may easily take many hours to complete even if run on modern hardware. CA algorithms do not solve the full hydraulic equations but employ simple operations of the CA rules; thus execution speeds could be orders of magnitude faster. Furthermore, CA algorithms are well suited for parallel execution on modern high performance hardware (Gibson et al., 2015), since the computation of a cell’s state depends only on the previous state of the local neighbouring cells and its previous state.

In CADDIES, simplified CA transition rules, GP methods and parallel computing tech-niques have been adopted to develop a series of models for simulating flow in sewer networks and on overland surface. Their applications on real world case studies show that the CADDIES1D one-dimensional sewer models were capable of simulating the flow dynamics with less computing time but good accuracy, as compared with a standard commercial soft-ware package (Austin et al., 2014). The improvement of efficiency is more significant in the CADDIES2D overland flow models, which performed eight times faster than the standard commercial software package on a large urban catchment (Guidolin et al., 2015). The suc-cessful developments have enabled the CADDIES models to be implemented for flood risk assessment at the continental scale, uncertainty analysis of flooding in urban areas due to pipe bursting, and near real-time flood modelling and forecasting.

The CADDIES project also includes a new software framework that provides developers with various tools that simplify development of parallel CA algorithms for pluvial flood simula-tion. The framework is composed of a graphical application and a CA application program-ming interface (API). The former is used to manage and visualise input and output data as well interfacing with GIS; thus the need for the developers to invest valuable time and resources on the creation of input and output code, GIS interface and visualisation tools is reduced. The CA API defines a standard set of methods, data structures and variables that can be used to develop parallel CA algorithms. The CA API can handle different implementations depending on the type of hardware and CA algorithm used. Thus it allows a developer to focus mainly on writing the rules of the CA algorithm, while the data management, parallelisation and execution of the algorithm is handled by the specific implementation of the CA API.

Further information of the CADDIES framework can be found at:
http://emps.exeter.ac.uk/engineering/research/cws/resources/caddies-framework/

A further machine learning technique, applied to urban flood modelling from multiple nodes in an urban drainage network simultaneously, is Artificial Neural Networks (ANNs) (Bishop, 1995). Early Warning Systems (EWS) have been created to predict flood depths, volumes or flow-rates from individual sewer nodes, such as manholes or Combined Sewer Overflows (CSO’s) and to classify flooding by levels of severity (Duncan et al., 2011, 2013). ANNs are a nature-inspired method based on the operation of neurons in the human brain or nervous system. In this application, 2-layered, fully-connected feedforward ANNs have been used. The network has a number of input signals, for example antecedent rainfall at a variety of timestep lags. It also has a single output for each sewer node for which flooding is to be predicted. During training using data from historical rainfall events, the multiplicative weights and biases at the inputs of each neuron are adjusted so as to minimise the error between the ANN’s outputs and the expected flood levels/volumes based on the historical or hydrodynamically simulated data. Once the ANN is trained, execution is extremely rapid since no iterative loops are involved in making flood predictions based on the new input rainfall data values. Being able to make predictions of flooding down to the sewer node level is extremely useful operationally, since it enables appropriate mitigation measures, such as sandbagging and/or alteration of sluice states to be carried out in advance of the flooding. Using this ANN approach, the RAPIDS (RAdar Pluvial flooding Identification for Drainage System) software tool has been developed and a variety of urban flooding case studies have been carried out using it (UKWIR, 2012).

More information on the RAPIDS system can be found at:
https://emps.exeter.ac.uk/engineering/research/cws/research/urban-drainage/rapids/
Acknowledgements

The authors would like to acknowledge the funding provided by the UK Engineering and Physical Sciences Research Council, grant EP/H015736/1 (Simplified Dual-Drainage Modelling for Flood Risk Assessment in Urban Areas) and UK Water Industry Research (UKWIR), grant SW01A201 (Real Time Machine Learning project).

Figure 1: Example of three CA grids with different attributes

Figure 2: Feedforward, Layered ANN Architecture

Figure 3: Flooding from surcharged manhole in street
References

UKWIR, 2012. The Use of Artificial Neural Networks (ANNs) in Modelling Sewerage Systems for Management in Real Time: Volume 1 - Main Report (12/SW/01/2) (Project Final Re-port No. 12/SW/01/2). UKWIR (UK Water Industry Research), London, UK.