NERIS Workshop 2017

17-19 May 2017
Lisbon, Portugal

Abstract Template

Please select one or more topics for your abstract

	RADIOLOGICAL IMPACT ASSESSMENTS
X	COUTERMEASURES AND COUTERMEASURE STRATEGIES IN EMERGENCY & RECOVERY, DECISION SUPPORT & DISASTER INFORMATICS
X	SETTING-UP A HOLISTIC FRAMEWORK FOR PREPAREDNESS FOR EMERGENCY RESPONSE & RECOVERY
	STAKEHOLDER INVOLVEMENT AND ENGAGEMENT IN EMERGENCY AND RECOVERY
X	UNCERTAINTY HANDLING ISSUES FOR EMERGENCY AND RECOVERY

Title

A Big Data Architecture for Learning-Based Source Term Estimation

Authors – Please underline corresponding author

Iraklis A. Klampanos (iaklampanos@iit.demokritos.gr), A. Davvetas,
C. Pappas, S. Andronopoulos, A. Ikonomopoulos and V. Karkaletsis
National Centre for Scientific Research "Demokritos", Aghia Paraskevi, Attiki, Greece

Text

Typically, after the detection of a radioactive substance of unknown origin in the atmosphere the source location is estimated via inverse modelling. Depending on various factors, such as the spatial resolution desired, traditional inverse modelling can be computationally time-consuming and therefore its application can be problematic when timing is critical. In a complementary presentation (S. Andronopoulos et al., “Towards Inverse Source Term Estimation using Big Data Technologies”), we will discuss a data-scientific approach to source term estimation, which allows us to perform the bulk of the processing prior to such an event taking place, therefore allowing for rapid estimation.

This presentation will be concerned with the Big Data Europe[1] (BDE) platform and application architecture which allows us to implement the algorithm above and perform experimental analyses. Part of the BDE project’s goal is to create a big-data application aggregator which is versatile and easy-enough to encourage data-related progress in a diverse set of areas of societal impact. The BDE platform provides an aggregation of widely-used and specialised tools used for big-data storage, processing and analysis. Its architecture is based on containerisation (the current implementation makes use of Docker[2] containers), which is a form of lightweight virtualisation and ensures deployability over a host of different platforms.

The design and implementation of our use-case involves two parallel pipelines of data transformation and control. The first is a batch processing dataflow, which processes and analyses weather data to create plume dispersions based on previously learnt weather patterns. As part of this use-case, we make use of a dockerised version of the WRF model[3]. The raw data are pre-processed to obtain coarse representations of the European domain in the NetCDF format[4]. In a machine-learning (ML) container, the coarse data are being analysed and clustered using several statistics-based approaches, before they are downscaled to achieve higher resolution versions of the weather patterns previously learnt. Using the DIPCOT dispersion model[5] (also available in JRODOS[6]), we pre-compute plume dispersions for each of the resulting weather pattern, before we store and index the results in a geospatially-aware database. Once these dispersions have been indexed, we can match newly observed weather patterns and retrieve potential dispersions efficiently.

The second pipeline is interactive and can be used once the analysis described above has been completed. We use a web-based platform for visualising linked geospatial data. The user selects one of a set of weathers previously unseen by the system (in an operational setting this would be the actual currently observed weather) and a few coordinates simulating locations where radioactivity has hypothetically been detected (or real detection locations, in an operational setting). Given this in-

---

[1] https://www.big-data-europe.eu
[2] https://www.docker.com
[3] http://www.wrf-model.org
[4] http://www.unidata.ucar.edu/software/netcdf/
[5] http://pandora.meng.auth.gr/mds/showlong.php?id=35
[6] https://resy5.iket.kit.edu/JRODOS/

formation the system estimates the potential release origin(s) by making use of the use-case's ML component, interrogating the pattern dispersions database as needed. The result is routed back to the user via the graphical user interface.

In this presentation, we will give more details about the BDE platform, the components used and implemented, as well as their performance, shortcomings and pointers for future work.