Anomaly detection in constrained devices

Laura Erhan, Antonio Liotta
Data Science Research Centre, University of Derby, Derby, UK

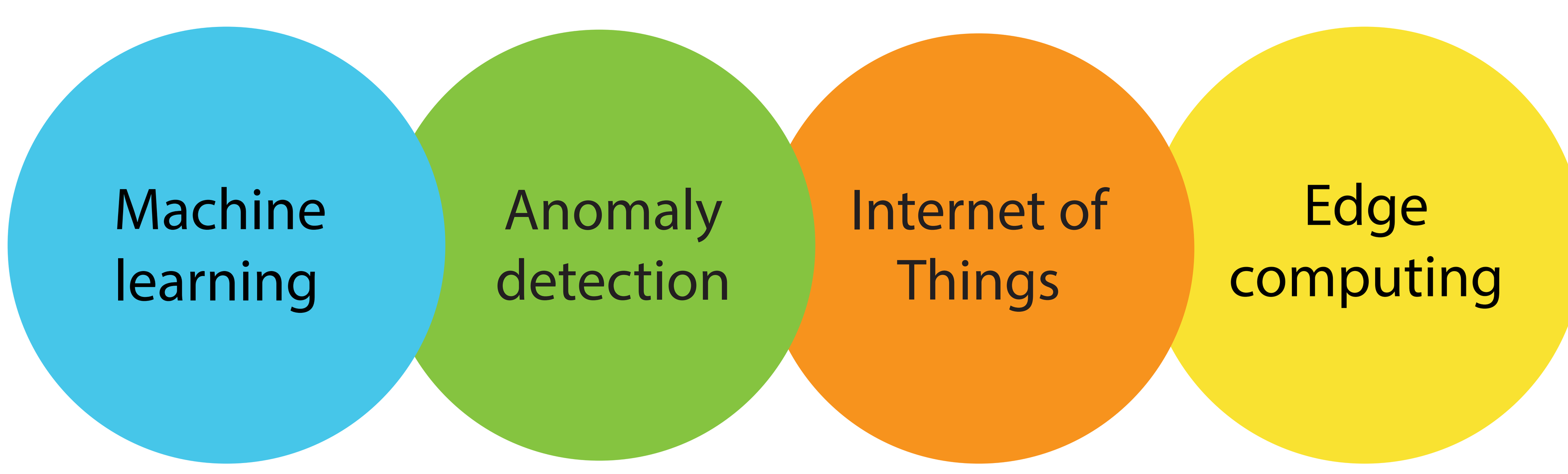
Contact: l.erhan@derby.ac.uk
Keywords: unsupervised learning, online learning, anomaly detection, edge computing, Internet of Things, constrained devices, in-node processing

UNIVERSITY OF DERBY

Summary

The proposed study aims to tackle the lack of an (online) learning framework which can be deployed at the edge. We are looking at machine learning based anomaly detection techniques in constrained devices (clusters of nodes at the edge of a system). The end goal is to achieve online (near real-time) anomaly detection in an IoT system, starting from the node devices of the system. By taking advantage of the communication between devices and exchanging information, the nodes could learn together from the data and from their environment and therefore, diminish the need for high computational power.

Background



- Number of IoT devices is fastly increasing
- Big data is being generated and needs to be analysed
- Current infrastructure (cloud-based) can be a bottleneck, especially in the case of streaming data
- Unsupervised anomaly detection in streaming data is still a challenge
- IoT devices are more powerful than they used to be (e.g. RPI)
- Processing can take place at the edge of the system, in the node devices
- Lots of applications require results online/near real-time
- Processing of data streams (time series) should happen in an unsupervised online fashion

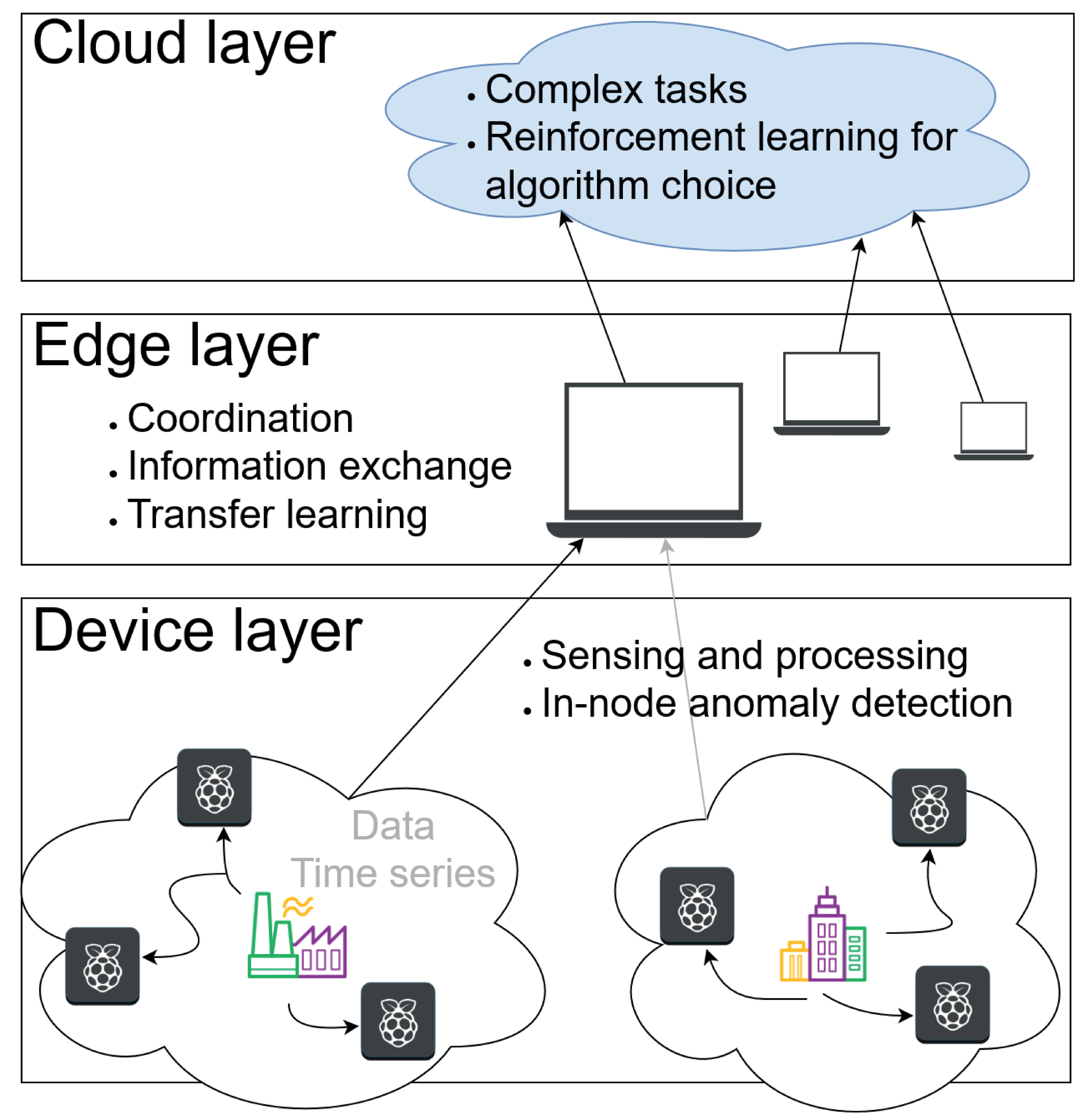
Objectives

- How do existing state of the art algorithms for unsupervised anomaly detection perform when the hardware resources are limited?
- How can these algorithms be modified in order to accommodate constrained devices?
- How can neighbourhood information be used in order to increase the performance of the system?
- What are the limitations of the proposed framework in terms of scale and processing power?
- Can reinforcement learning be successfully used in order to select the appropriate algorithm to detect anomalies in a given scenario?
- How does the fully integrated (cloud – edge – device) proposed anomaly detection framework perform in a real-case scenario?

The approach chosen for the project is bottom-up. This means that the first step includes looking at the node device in order to assess both its capabilities and its limits. The chosen node is a Raspberry PI, a mini computer. For this assessment, a group of unsupervised real-time anomaly detection algorithms was chosen, namely the algorithms part of NAB which is an anomaly benchmark which also comprises of multiple real-life and artificial time series datasets.

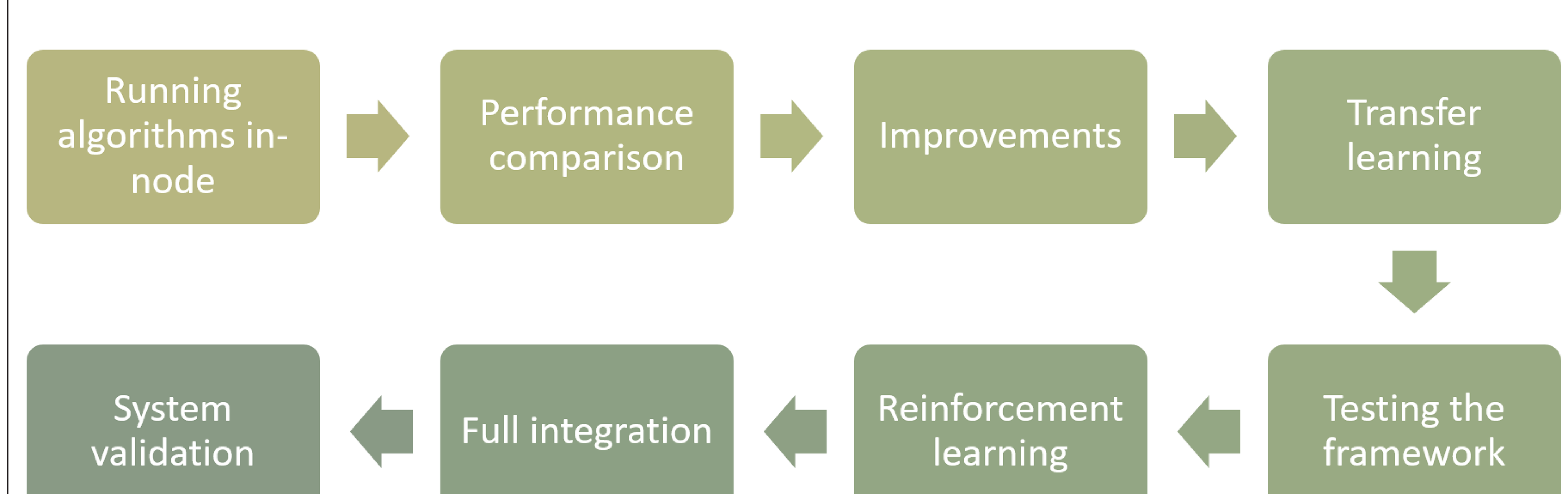
Proposed solution

We propose an online anomaly detection framework for clusters of IoT nodes.



Next steps

In order to test the feasability of running the algorithms in a constrained device such as the RPI, the NAB benchmark was ran on different configurations of virtual machines (varying RAM memory). The approach validated that running with limited memory can be achieved; however, it also highlighted certain problems that need to be fixed first (such as modifying the algorithm to avoid bad allocation errors, library comptatibility issues etc.). The next steps to be taken are highlighted below:



References

S. Ahmad, A. Lavin, S. Purdy, and Z. Agha, 'Unsupervised real-time anomaly detection for streaming data', Neurocomputing, vol. 262, pp. 134–147, Nov. 2017.
H. H. W. J. Bosman, G. Iacca, A. Tejada, H. J. Wörtche, and A. Liotta, 'Ensembles of incremental learners to detect anomalies in ad hoc sensor networks', Ad Hoc Networks, vol. 35, pp. 14–36, Dec. 2015.
H. H. Bosman, G. Iacca, A. Tejada, H. J. Wörtche, and A. Liotta, 'Spatial anomaly detection in sensor networks using neighborhood information', Information Fusion, vol. 33, pp. 41–56, Jan. 2017.
M. Lavassani, S. Forsström, U. Jennehag, and T. Zhang, 'Combining Fog Computing with Sensor Mote Machine Learning for Industrial IoT', Sensors, vol. 18, no. 5, p. 1532, May 2018.