# A Bi-objective Hyper-heuristic Support Vector Machines for Big Data Cyber-Security

**Abstract:**

Cyber security in the context of big data is known to be a critical problem and presents a great challenge to the research community. Machine learning algorithms have been suggested as candidates for handling big data security problems. Among these algorithms, support vector machines (SVMs) have achieved remarkable success on various classification problems. However, to establish an effective SVM, the user needs to define the proper SVM configuration in advance, which is a challenging task that requires expert knowledge and a large amount of manual effort for trial and error. In this work, we formulate the SVM configuration process as a bi-objective optimisation problem in which accuracy and model complexity are considered as two conflicting objectives. We propose a novel hyperheuristic framework for bi-objective optimisation that is independent of the problem domain. This is the first time that a hyper-heuristic has been developed for this problem. The proposed hyper-heuristic framework consists of a high-level strategy and low-level heuristics. The high-level strategy uses the search performance to control the selection of which low-level heuristic should be used to generate a new SVM configuration. The low-level heuristics each use different rules to effectively explore the SVM configuration search space. To address bi-objective optimisation, the proposed framework adaptively integrates the strengths of decomposition- and Pareto-based approaches to approximate the Pareto set of SVM configurations. The effectiveness of the proposed framework has been evaluated on two cyber security problems: Microsoft malware big data classification and anomaly intrusion detection. The obtained results demonstrate that the proposed framework is very effective, if not superior, compared with its counterparts and other algorithms.

**Index Terms**—Hyper-heuristics, Big data, Cyber security, Optimisation.