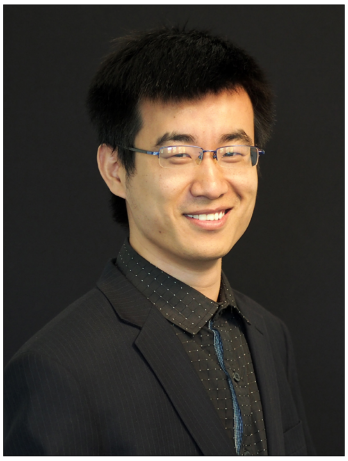
Capital Markets CRC Limited

Data mining researchers use innovative techniques to build robust classifier



Fei Wang
CMCRC PhD candidate
U. of Sydney

Prof Sanjay Chawla
CMCRC Research Leader
U. of Sydney

Researcher discovers that a combination of adversarial learning and sparse modelling techniques improves the performance of an email/spam classifier.

Classifiers are widely used in many computer-based applications to the benefit of virtually all computer users. An important application of these classifiers is email filtering which decides whether an incoming email is a proper email (ham) or spam. Research by Fei Wang and research supervisor Professor Sanjay Chawla have combined sparse modelling and adversarial learning techniques to produce a spam classifier that outperforms the current model used in email applications. The technique has wide applicability and is also used in Wang's extensive analysis of health insurance claims data for CMCIS.

Most classifier applications are trained by giving them a certain amount of labelled data (training data) which have basically the same characteristics as the data the classifier is going to classify (test data). However, there are a couple of major obstacles that make the design of the perfect classifier with longevity rather difficult. Firstly, there are situations where the training data may have different characteristics from the test data, for example peoples buying patterns may be influenced by seasons, more ice-creams are sold in Summer than Winter for instance (known as *concept drift*). Generally, the characteristics of the test data never completely match that of the training data. Secondly, there may be some kind of adversarial influence which may affect the performance of the classifier over time. An example of this is the spammer who tries his best to formulate an email to circumvent a classifier. Over time, the spammer will find more and more ways to beat the classifier making it obsolete over time.

Wang's research combines adversarial learning with sparse modelling techniques (this discovers predictive patterns in data) into a repeated game to make the research realistic with the real world. The role of sparse techniques is to model the scenario that spammers (and those working to prevent spam) have limited budgets. The traditional approach to keep the classifiers updated is to repeatedly build the classifier in the face of changing data. In Wang's research however, ideas from game theory are used to characterise an equilibrium, which has the side effect of creating new training data. The new training data, which in some sense, anticipates future adversarial behaviour is then used to build the classifier. Fei Wang shows that this approach is both robust and economical compared to previous approaches.

The Capital Markets Cooperative Research Centre is a world-leading research organisation that provides thought leadership and break-through technology solutions for capital and insurance markets (www.cmcrc.com). CMCIS (www.cmc-is.com) is a spinoff company and partner to the CMCRC which delivers intelligence to the health and accident compensation insurance industries.

In partnership with CMC Insurance Solutions