The detection of malicious executables (malware) is a well known problem. Anti-malware software is typically signature based, and only malicious attacks containing those known signatures can be detected. This is problematic because new malware is appearing extremely rapidly. This threatens to overwhelm signature-based approaches. In this paper, we propose a novel approach to detect malicious executables by using a combination of techniques from bioinformatics, data mining and information retrieval. This method is able to identify new malware related to threats already in its database. Using relatively small training sets our technique is able to achieve over 90% accuracy of detection with a false positive rate smaller than 5%.