

Information and Computer Technologies

Research on characteristic parameters mining and clustering of unknown protocols bitstreams

Yang Wu, Tao Wang, Jin-dong Li

Computer Modelling & New Technologies 2015 19(2B) 7-15

Characteristic parameters mining of unknown protocol bitstreams and parameters optimizing of clustering algorithm are the foundations of unknown protocol bitstreams analyzing. The parameters such as the bit frequency, runs and bit frequency within a block are defined according to the frequency of zero and one, frequency of sequential zero and one, bit frequency within a block. As the parameter of bit frequency within a block is sensitive to the block length, an optimal block length selection algorithm is proposed based on the principle of variance. In order to select effective initial clustering centers for division clustering algorithms such as the k-means algorithm, an initial clustering centers selection algorithm is proposed based on the peak value of sample density for each dimension. In order to select the optimal clustering number, a function of clustering quality evaluation is given by the sample density in cluster and cluster density. Taking the bitstreams of HTTP, DNS, ICMP, TELNET and UDP datasets as the unknown protocols bitstreams, the experimental results not only verified the effectiveness of the proposed algorithms but also point out the necessity of mining more effective parameters.

Keywords: Unknown protocol, bitstreams, clustering, characteristic parameter, bit frequency within a block

Cache Pre-fetching system based on data mining on Web

Yongbiao Li

Computer Modelling & New Technologies 2015 19(2B) 16-21

From 20th century 90's to now on, Internet and data mining techniques had developed rapidly and became mature, kinds of application on Web data mining had been proposed to the market. In this paper, we would first introduce the development of cache Pre-fetching technique, and then present a cache pre-fetching System model based on Web data mining, details of each implementation would follow. Our aim was to enhance caching effectiveness, and network accessing speed. Such technique could be applied in personnel, educational, and official information managing system in database of educational scope. Accessing speed of educational information system for numerous teachers and students, benefit high school personnel management, and also the effective scientific structuralize educational management.

Keywords: Web data mining, sequential mining, cache pre-fetching system

High resolution remote sensing image classification based on particle swarm optimization and support vector machine

Buyi Li, Chongjing Deng, Shuang Li

Computer Modelling & New Technologies 2015 19(2B) 22-26

Many algorithms have been developed for image classification and support vector machine (SVM) is a kind of supervised classification that has been widely used recently. However, the accuracy of a SVM classifier heavily depends on the selection of a right kernel model and appropriate parameter. In this paper, a comparative analysis of the impact of four kernels (linear kernel, polynomial kernel, radial basis function kernel and sigmoid kernel) on the accuracy of SVM classifiers is conducted. Moreover, the Particle Swarm Optimization (PSO) is used to search for the optimum parameters for each kernel function in order to improve the classification accuracy of SVM classifiers. Our experiments for optimizing the kernel function parameters and assessing the robustness of SVM classifiers were carried out with classifications of QuickBird-2 images over Wuhan, China for monitoring urban land cover/land use information. The experimental results indicate that the polynomial kernel outperforms the other kernels in classifying high resolution remote sensing image. The sigmoid kernel performs worse than any other kernels. Our findings also suggest that selected parameter by PSO will improve the classification accuracy, especially for radial basis function kernel.

Keywords: high resolution remote sensing image, support vector machine classification, parameter optimization, particle swarm optimization

Pattern recognition systems in the problems of automatic person identification using the passport data

Y Amirgaliyev, R Yunussov

Computer Modelling & New Technologies 2015 19(2B) 27-30

The work describes the implementation of modern technology for remote sensing and data processing in the area of human activities concerned to the security provision, based on usage of pattern recognition algorithms and neural networks. The Republic of Kazakhstan State Identities and Passports were used as the basis; The ICAO 9303 MRZ Standard was used. Obtained stable recognition model for identification of known passport types, and MRZ section decoding.

Keywords: neural networks, pattern recognition, raster data, image processing