

Performance Evaluation of Local Area Network based on Support Vector Machine

Qi Liu^{1, a}, Yin Liu^{1, b}, Yiyong Lin^{2, c}, Ling He^{1, d} and Yunzhi Huang^{1, e}

¹school of Electrical Engineering and Information, Sichuan University, Chengdu, China

²Xichang Satellite Launch Center, Xichang, China

^aliuqi@scu.edu.cn, ^b79981276@qq.com, ^c799729599@qq.com, ^dling.he@scu.edu.cn, ^ehuang_yunzhi@scu.edu.cn

Keywords: IP performance evaluation; support vector machine; k-fold cross validation; local area network

Abstract. This work presents an IP performance evaluation method, based on the Support Vector Machine (SVM). In this work, eight network parameters are collected: CPU utilization of switchboard, utilization of memory, network link-off, delay, delay jitter, bandwidth, bit rate of transmission and bit rate of reception. The performance of the local area network is classified into three grades: excellent, good and failed. In this work, the collected network parameters are processed using SVM classifier. The average classification accuracy is achieved using k-fold cross validation method. The experiment results indicate that the classification accuracy of the proposed method is over 90%. This proposed evaluation system could be applied to a real-time network performance evaluation application effectively.

Introduction

The Internet technology has developed rapidly in the present days. Currently, many researches have been done to study the network reliability, stability, efficiency and other properties of network. The accurate IP performance evaluation of local area network has a significant meaning. Through the analysis of the network core equipment's handling capacity, network bandwidth utilization and network load, the overall network performance could be improved. Moreover, it helps to provide a reliable basis for the planning and adjusting of networks.

Currently, researchers has studied deeply on the methods of network performance evaluation. The international organization TETF (Internet Engineering Task Force) and ITU-T have released the network performance evaluation standard. However, these documents have not provided the specific performance evaluation model [1, 2]. Weike Que. et al [3] have proposed a comprehensive IP network performance assessment model. This method combines the concept of network health and the health of router. It makes an overall evaluation of the network by measuring the basic collected data. However, the multivariate index weight is difficult to determine in the process of the network operation quality evaluation. Qingpu CAO et al. [4] have used a projection pursuit method and hybrid genetic algorithm to determine the indicator weights objectively in the process of network performance evaluation. S. Sathiya Keerthi et al. [5, 6] evaluate the network performance using fuzzy mathematics methods. Li et al. [7] extract the wavelet based features and use neural network to achieve the network performance evaluation.

Combined with pattern recognition technology, this work presents an IP performance evaluation algorithm based on the SVM classifier. The software platform is Visual Studio 2012. Eight network parameters are collected in this work: CPU utilization of switchboard, utilization of memory, network link-off, delay, delay jitter, bandwidth, bit rate of transmission, bit rate of reception. Combined with multi-classification SVM, the system evaluate the performance of local area network in three grades: excellent, good and failed.

Methods

I. Parameters of Local Area Network

For different LANs, the parameters which reflect the network's performance may differ. For example, for remote medical communication network, the network delay and bandwidth are the most significant indicators in the network performance evaluation. While for some public service networks, the network security and throughput are important. The basic principles of choosing the most reasonable parameters for the network performance evaluation are: streamline and testability.

In this local area network, the parameters of switchboards and communication links are important for network performance evaluation. The CUP utilization, memory utilization of switchboard are chosen as the evaluation parameters. And the network link-off, delay, delay jitter, bandwidth, bit rate of transmission, bit rate of reception of the communication links are chosen as the links' parameters.

II. Support Vector Machine

A. Theory of Support Vector Machine. The basic idea of SVM is to find the optimal classification plane for two types of samples, which causes the maximum classification distance of these two kinds of samples. SVM is established on the basis of statistical learning's VC dimension theory, as well as structural risk minimization theory. It avoids the local optimal solution and cleverly overcomes the "curse of dimensionality". Especially, it has a great of advantages in solving problems with small samples, or with nonlinear and high dimension input space [8].

SVM is firstly used to deal with the dichotomous linearly separable problem, for linearly separable sample collection $(x_i, y_i), i=1 \dots, n, x \in \mathbb{R}^d, y \in \{+1, -1\}$, it satisfies Eq. 1:

$$y_i[(w \cdot x_i) + b] - 1 \geq 0 \quad (1)$$

One of core ideas of SVM is to maximize the classification distance. Lagrange optimization method can translate the optimal classification problem to a dual problem. With unique solution, the optimal classification function is:

$$f(x) = \text{sgn}\{(w \cdot x_i) + b\} = \text{sgn}\left\{\sum \alpha_i^* y_i (x_i \cdot x) + b^*\right\} . \quad (2)$$

Among the solutions of the above equation, there are only part of α_i which is not equal to zero. The corresponding samples are support vectors. In fact, the above equations only sums the support vectors. b^* is the classification threshold. In the case of linear inseparable, a relaxation term $\xi_i \geq 0$ is added, and Eq.2 becomes to Eq.3:

$$y_i[(w \cdot x_i) + b] - 1 + \xi_i \geq 0, i = 1, \dots, n . \quad (3)$$

By adding a penalty parameter C and changing the target into calculating the minimum value of Eq.4, a generalized optimal classification surface is obtained as:

$$(w, \xi) = \frac{1}{2} \|w\|^2 + C \left[\sum_{i=1}^n \xi_i \right] . \quad (4)$$

The nonlinear problem could be transformed into a linear problem in a high-dimension space. A optimal classification surface could be achieved in the high-dimension space. The linear classification could be performed by selecting a proper kernel function. And the classification equation becomes as Eq.5:

$$f(x) = \text{sgn}\{\sum \alpha_i^* y_i k(x_i \cdot x)\} + b^* . \quad (5)$$

The SVM achieves the nonlinear transformation by using the inner product function, which transforms the input space into a high-dimension space, and calculates the generalized optimal classification surface in the high-dimensional space. The inner product functions are also called as kernel functions. So far, there are three types of kernel functions [9]:

(1) Polynomial inner product function (Eq.6):

$$k(x, x_i) = [\gamma(x \cdot x_i) + C]^q . \quad (6)$$

(2) Radial basis inner product function (Eq.7):

$$k(x, x_i) = \exp(-\gamma \|x - x_i\|^2) . \quad (7)$$

(3) Sigmoid inner product function (Eq.8):

$$k(x, x_i) = \tanh[\gamma(x \cdot x_i) + C] . \quad (8)$$

This work chooses radial basis inner product function RBF as the kernel function of SVM. The fault classification penalty factor C and RBF kernel function's control factor γ are the two important parameters for RBF inner product function.

B. Multi-classification SVM. Multiple SVM binary classifier [10] combinations are usually used to deal with the multi-classification problems. The commonly applied combination methods include: one-to-one algorithm, one-to-many algorithm, directed acyclic graph algorithm, hierarchical algorithm, error correction coding algorithm, etc.. In this work, the one-to-many combination method is adopted. Fig.1 indicates the schematic diagram of a one-to-many classification problem.

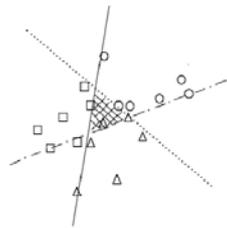


Fig. 1. One-to-many classification schematic diagram

For the database with N types, one-to-many algorithm utilizes SVM to construct N binary sub-classifiers. The construction method is as follows: the i-th SVM binary sub-classifier considers the data with class i as a positive category, and the remaining N-1 classes data as a negative category. Therefore, the N types classification problem transforms to the N SVM binary sub-classification.

Fig.2 describes the structure of a multi-classification classifier group, which is applied in this work.



Fig. 2. The structure of a multi-classification SVM classifier group

Experiments and Results

In this work, the local area network contains 89 switchboards and 168 data links. This system is based on Visual Studio 2012 development platform. Combined with the SNMP protocol, this system achieves the real-time parameters collection of switchboards and communication links. The network parameters include CPU utilization of switchboard, memory utilization, network link-off, delay, delay jitter, bandwidth, bit rate of transmission and bit rate of reception. For every two seconds, all these parameters are collected from all the switchboards and data links in the whole local area network. The SVM classifier is applied to achieve the network performance evaluation into three grades: excellent, good and failed. A total of 3000 groups datasets are used in this work. The 10-fold cross validation algorithm is used to obtain the classification accuracy. The correct classification rates for three grades of network performance is 92%. And the running time of the system for each time evaluation is less than 0.5ms.

Discussions and Conclusions

This passage presents an IP performance evaluation method based on a multi-classification SVM classifier. The accuracy of this system reaches up to 92%, and the running time of this system is less than 0.5ms. This system could be applied as a real-time application in the local area network.

References

- [1] ITU-TY. 1540-2007 internet protocol data communication service-IP packet transfer and availability performance parameters[S]. ITU, 2007.
- [2] SIEPHAN E. IETF RFC 4148 IP performance metrics (IPPM) metrics registry[S]. Internet engineering task force, 2005.
- [3] Weike. QUE, Guoqing ZHANG, Zhenghao. WEI Comprehensive evaluation model of IP network [J]. Computer engineering, 2008, 34(8): 99-101.
- [4] Qingpu CAO. Evaluation of network performance using genetic projection pursuit model weighted method [J]. Computer simulation, 2010, 27(10): 103-106, 262.
- [5] Causeur, D. "Factor Analysis for Multiple Testing (FAMT): An R Package for Large-Scale significance testing under Dependence." Journal of statistical software 40.14 (2011): 1-19.
- [6] Neural Computation. Jul2003, Vol. 15 Issue 7, p1667-1689. 23p.
- [7] H. Li, D. Doermann, and O. Kia, "Automatic Text Detection and Tracking in Digital Video," IEEE Trans. Image Processing, vol. 9, no. 1, pp. 147-156, 2000.
- [8] Gang Sun .Intrusion detection system of SVM [D].Beijing: Beijing university of Posts and Telecommunications, 2007.
- [9] Kim, K. "Kerbel principal component analysis for texture classification." IEEE signal processing letters 8.2 (2001): 39-41.
- [10] Lai C Y. A Novel Approach to Multi-classifier Based on Multiple feature Sets with SVM for Network Intrusion Detection [D]. Taiwan: Chung Hua University, 2007.