ISSN : 0974 - 7435 Volume 10 Issue 24

2014

BioTechnology
An Indian Journal

FULL PAPER

BTAIJ, 10(24), 2014 [15077-15081]

Research on remote vehicle intelligent diagnosis based on KNN

Miu Kehua, Li xiaokun*
Xiamen University Pattern Institute, 361005, Xiamen, (CHINA)
E-mail : Haoyuan101@qq.com

ABSTRACT

This paper provides a remote vehicle diagnosis system, which is designed to locate the specific time when an occasional malfunction happened from the abundant vehicle's ECU data flow. The system has been designed with an ability to learn by itself, using the wrong cases to retrain the classifier and raise system diagnosis rate. Through studying the occasional low-speed flameout, we come to a conclusion that 83.3% diagnosis rate and nanosecond-class diagnosis efficiency can totally meet requirement.

KEYWORDS

Remote diagnosis; Vehicle system; KNN; Self-learning; Occasional malfunction.

INTRODUCTION

Occasional vehicle malfunctions, such as low-speed flameout, difficulty in starting, and four-cylinder fire lacking, etc., not only bring great distress to vehicle users, but also cause security risks. Remote vehicle diagnosis system which remotely read and store vehicle ECU running data and remotely analysis the stored data can trace and solve the problem of occasional malfunctions [1]. KNN is a mature and simple method in pattern classification algorithms, which has certain advantages in both training and recognition stages.

Instruction on remote vehicle diagnosis

A remote vehicle diagnosis system is composed of remote diagnostic equipment, client software, and a server [2]. The system working principle [3] is shown in Fig 1.

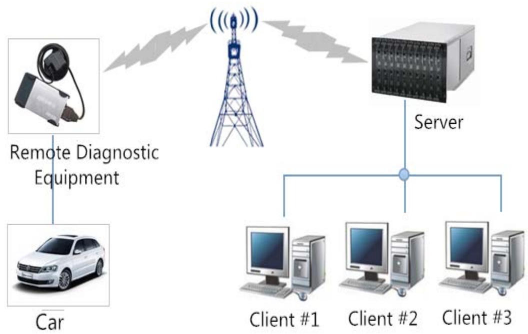

Figure 1 : Basic principles of the remote diagnosis system

The remote diagnostic equipment is installed on the vehicle. It accesses the control unit of each part of the vehicle and acquires data. The data is transmitted to server in real time. The authorized users can use client software to remotely operate the diagnostic equipment and explore, analyze as well as download the data flow and fault codes of each part of vehicles during running. The server receives and stores vehicle running data. Meanwhile, the server tracks and locates the time of occasional malfunctions in the massively stored vehicle data. The malfunction data at the located time are returned to professional and technical personnel to help them accurately analyze and solve the occasional malfunctions[4,5].

Brief introduction on KNN

KNN is a kind of simple yet effective method [6]. Let a reference sample set be Dn={X1,...,Xn}. The class membership of each member Xi in the set is known. For a test sample X, its class depends on the votes of K closest samples to it in set Dn.

Because samples near boundaries between classes interlace, KNN method often misclassifies such samples. Clip method can be used to cleanup boundaries between classes and remove fuzzy samples, making the class boundaries clearer [7].

The system has the ability of self-learning, therefore the number of reference samples will increase continuously. To avoid calculating the distance between test samples and each reference sample, the C-means clustering method is combined to establish a tree searching structure (a binary tree) in the reference sample set, which improves the diagnosis classification efficiency of the system[8].

RESULT AND DISSCUSS

The remote vehicle diagnosis system established in this study is applied to Shanghai Volkswagen Automotives (SVA). There are more than 1,000 vehicles for daily simultaneous online diagnosis. For each fault vehicle, the system simultaneously collects dozens of measurement values of running vehicle data in several ECUs by millisecond intervals. The stored vehicle running data are very huge in amount, but the occasional failure frequency is very low, even several months once. Therefore, manual analyzing these data in pieces will cost a lot of manpower and material resources. It is necessary to implement a system which can automatically recognize the failure points.

To realize the system function of automatic analyzing and recognizing failures, SVA remote intelligent diagnosis system uses an intelligent classifier with a two-stage learning mechanism. In the first stage, the intelligent diagnosis classifier is initially trained: The malfunction data are collected and used as training samples. The initial reference samples and the parameter value of K of the KNN classifier are obtained. In the second stage, the vehicles are diagnosed online and the diagnosis classifier learns by itself: The initially trained classifier is connected to the system to analyze data in real time. The recognition results are judged by professional and technical personnel. Then the misclassified data are used as training samples to retrain the classifier, which realizes the online learning function of the system. The flowchart of the established system is shown in Fig 2. “Comment 1” represents system online diagnosis.

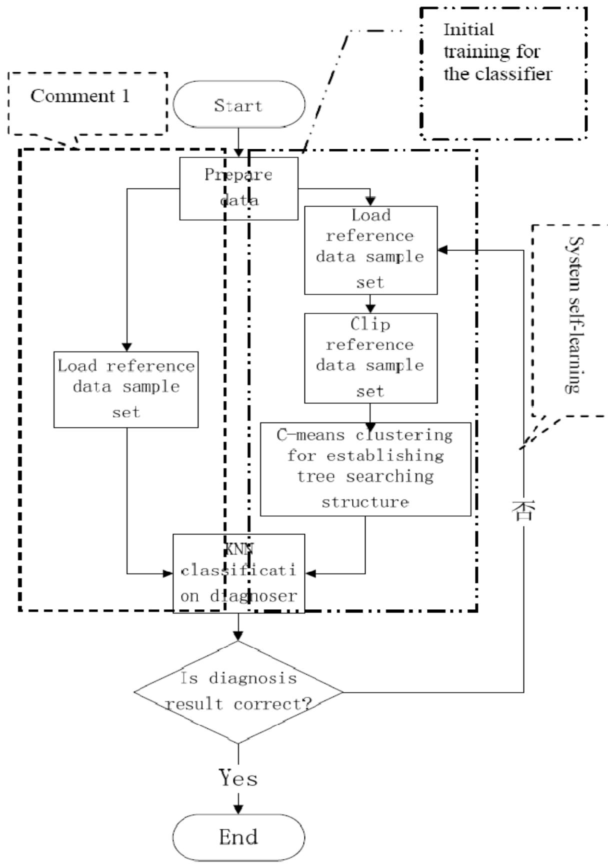

Figure 2 : The flowchart of the remote intelligent diagnosis system

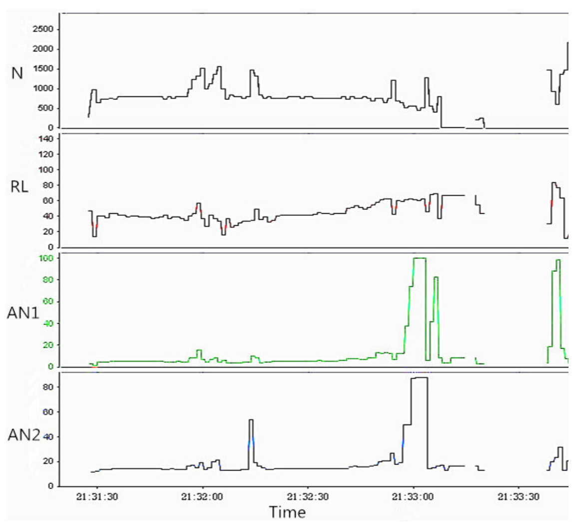

Figure 3 : Waveforms of the parameters related to occasional low-speed flameout malfunctions

Diagnosing vehicle low-speed flameout malfunctions

We take low-speed flameout as an example. The relative parameters include engine speed (N), relative load (RL), throttle valve angle (AN1), accelerator pedal angle (AN2), intake air amount (GI), average fuel injection time (T), fuel pressure (P1) and boost pressure (P2). Around the occurrence of low-speed flameout malfunction, the waveforms of the engine speed, the relative load and the throttle valve angle are shown in Fig 3.

To make the classifier compatible, we sample data of three fault vehicles before and after the malfunctions and extract 100 data sets for training samples. Meanwhile, there are only 2 kinds of decision attributes D, where “1” stands for low-speed flameout malfunction, while “0” represents normal. The data are tabulated in Table 1. Moreover, data of another fault vehicle around its malfunction are extracted for test samples, totaling 30 data sets, which are shown in Table 2.

TABLE 1 : Training samples for low-speed flameout malfunction

Vehicl e No.	RL	T	N	GI	AN1	AN2	P1	P2	D
1	68.421	5.865	680.0	8.25	41.176	8.984	41.3	1310.0	0
1	67.669	6.63	0.0	0.777	3.921	7.031	40.0	1180.0	1
…									
2	69.172	5.61	840.0	12.277	82.352	12.109	41.85	590.0	0
2	67.669	6.63	0.0	0.722	3.921	7.031	40.0	950.0	1
…									
3	67.669	6.63	560.0	4.861	8.627	7.031	40.0	1180.0	0
3	67.669	6.63	0.0	0.694	8.235	7.031	40.0	950.0	1

TABLE 2 : Test samples for low-speed flameout malfunction

Vehicle No.	RL	T	N	GI	AN1	AN2	P1	P2	D
1	69.172	5.01	833.7	12.277	82.352	11.109	41.85	590.0	0
2	65.421	5.844	680.0	8.25	41.176	8.914	41.3	1310.0	0
…									
29	66.328	6.63	0.0	0.722	3.921	7.031	40.0	950.0	1
30	67.669	6.63	0.0	0.777	3.921	7.031	40.0	1180.0	1

Analysis of initial training results

There are 100 training samples, 94 of which remain after clipping. A three-layer tree search structure is established. Therefore, about 26 distance calculation operations and one sorting operation are needed for each test sample. As a result, the diagnosis time is nanosecond-class, which is of high-efficiency. Table 3 shows the diagnosis results for the test samples.

Table 3 : Diagnosis results under different k values

K values	1	3	5	7
Diagnosis accuracy	70.2%	73.4%	83.3%	90.4%

It can be seen from Table 3 that when 94 training samples are used for diagnosing 30 test samples, the recognition rate is significantly improved with the increasing of the associated number nearest neighbors k, However the overall diagnosis rate is still not high. Analyzing the diagnosis result, it is clear that the misdiagnosed samples are mainly marginal samples, while the diagnosis for samples which can be clearly classified is relatively accurate. In real applications, a diagnosis rate of 83.3% is generally adequate to meet demands. Therefore in real systems, k is set to 5.

Effectiveness analysis of online self-learning

Because the system is capable of self-learning, it can automatically add misdiagnosed samples to its training samples, increasing the coverage of the training samples. As a result, the recognition ability of the system for marginal samples is improved, and the diagnosis ability of the system will be improved constantly. With the running of the system, the reference sample number of the classifier will increase. The variation of the consequential diagnosis rate is shown in Fig 4. It can be seen that with the adding of reference samples, the system diagnosis rate significantly increases. When the number of reference samples reaches 350, the system diagnosis rate for K=3 reaches over 90%, while the system diagnosis rate for K=5 reaches as high as 94.4%.

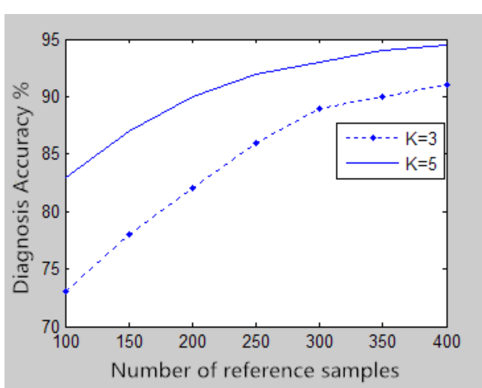

Figure 4 : Effectiveness of system self-learning

CONCLUSIONS

In this study, KNN classification algorithm is used to remotely diagnosis vehicle malfunctions. The following conclusions can be achieved:

1. KNN classification algorithm performs well in remote intelligent vehicle diagnosis, and it meets the demand on diagnosis efficiency and accuracy.

2. The self-learning mechanism is used to establish and improve the reference samples of the intelligent diagnosis system, avoiding the blindness and one-sidedness of extracting samples, which is a very effective method.

ACKNOWLEDGEMENT

Express our thanks to Shanghai Volkswagen Automotive Co., Ltd After-sales service Division for technical assistance and financial support.

REFERENCES

[1] Song You, Mark Krage, and Laci Jalics, "Overview of Remote Diagnosis and Maintenance for Automotive Systems", SAE World Congress, Detroit, Michigan, April 2005.

[2] Taie M A, Diab M, ElHelw M. Remote prognosis, diagnosis and maintenance for automotive architecture based on least squares support vector machine and multiple classifiers//Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2012 4th International Congress on. IEEE, 2012: 128-134.

[3] Kai C, Fuwu Y. A Remote Vehicle Diagnosis and Maintenance System Based on Internet//Electrical and Control Engineering (ICECE), 2010 International Conference on. IEEE, 2010: 228-231.

[4] Yan wufu, Cao ka, Hu jie,et al. Research on remote vehicle diagnosis data acquisition technology based on internet and 3G., 2013, 5: 017.

[5] Lin J, Chen S C, Shih Y T, et al. A study on remote on-line diagnostic system for vehicles by integrating the technology of OBD, GPS, and 3G. World Academy of Science, Engineering and Technology, 2009, 56: 435-441.

[6] Weinberger K Q, Saul L K. Distance metric learning for large margin nearest neighbor classification. The Journal of Machine Learning Research, 2009, 10: 207-244

[7] Weinberger K Q, Blitzer J, Saul L K. Distance metric learning for large margin nearest neighbor classification//Advances in neural information processing systems. 2005: 1473-1480.

[8] P. Y. Simard, Y. LeCun, and J. Decker. Efficient pattern recognition using a new transformation