RESEARCH ARTICLE OPEN ACCESS

A Survey on Vehicle Health Monitoring and Prediction System

T. Gayathri
Assistant Professor
Department of Computer Science Engineering
New Horizon College of Engineering
Bangalore, Karnataka - India.

ABSTRACT

Vehicle health monitoring utilizes the sensor data on the vehicle, mine the data and predicts the health of the car. It provides an idea on when the vehicle would need maintenance. The model also can warn when driver dozes off or in case of drunken driving. The purpose of this research paper is to provide an overview of the existing Vehicle Health Monitoring System. It discusses the different ways in which Vehicle Health can be monitored. Vehicle health monitoring can be distributed or on-board data mining. Computation or mining in vehicle are restricted by less memory and processor capacity, still an onboard mining is more advantageous than fully distributed datamining system. This paper survey the existing model for computationally efficient Onboard Vehicle Health monitoring system.

Keywords :— Vehicle Health Monitoring, data mining, prediction

I. INTRODUCTION

Vehicle come with 60-100 sensors which help in their effective monitoring. Digital electronic systems make it easy to constantly monitor vital engine parameters like oil pressure, coolant temperature and exhaust emissions and report back to the driver when something is amiss. In this paper, we try to exploit these real time data in predicting the performance deterioration in vehicles.

A vehicle can undergo an accident or break down anytime in its run. When the sensor data is monitored and mining is performed it would be possible to predict the good health or bad health of a vehicle.

Certain organization have huge fleet of trucks. Regular maintenance of vehicles in such fleet is an important part of supply chain management. There are also many Fleet management companies which allows companies which rely on transportation in business to remove or minimize the risks associated with vehicle investment, improving efficiency, productivity. For this purpose these companies collect the performance data, study the data offline and send the vehicles to service if necessary.

II. LITERATURE SURVEY

Many of the existing fleet management system uses one of the two approaches: One of them is discussed in [6]. The sensor data from the vehicle is retrieved using a control module and the parameters are uploaded to the fleet database. In the fleet database this data are analysed and the performance parameters of a vehicle to be diagnosed is analysed with the normal operation of the corresponding parameters to determine whether the vehicle to be diagnosed operates outside or whether it is a normal operation. The disadvantage of this approach bandwidth available is limited, it is not possible or efficient to transmit the performance parameters to the fleet database.

VEDAS [1] works on the concept of Ubiquitous Data Mining. An in-device, real-time mining of data on a ubiquitous computing environment in accordance to the environment's requirements by considering resource constraints of the device, exploiting context information, behaving autonomously and applying special privacy preserving methods. In VEDAS and MineFleet, Partial on board datamining is done.

This is the second approach. VEDAS is a mobile data stream mining environment. VEDAS analyses the data produced by the various sensors on the vehicle. It continuously monitors data stream generated by a moving vehicle using an on board computing device, identifies emerging pattern and report these patterns to central control server using low wireless network connection if necessary.

The data mining approach here is suited for small devices and thus its designed to be power efficient and uses less bandwidth.[1] has discussed experimental setup for displaying the energy consumption. The result show that on board analysis is better than broadcasting the sensor data and providing remote data mining.

The parameters that are analyses in Engine load(%), Engine torque, Intake Air temperature(°F), Engine Coolant Temperature (°F), Barometric Pressure, Mass Air Flow Sensor 1(MAF) (lbs/min), Start Up Intake Air Temperature (°F), Vehicle Speed (Miles/Hour), Throttle Position Sensor (%) , Throttle Position Sensor (degree), Start Up Engine Coolant Temp. (°F) and odometer.

Minefleet [2] an implementation model of VEDAS. This model for mobile on board data mining process was done in the year 2010 by the same author. Here the significant improvement which was made by them was the change in PCA (Principle Component Analysis) creation. FCM was used to create an approximate eigen vector, rather than using the 1naive approach.

III. FOCUS OF STUDY

The device that perform analysis on vehicles are PDA. The data mining task are usually compute intense. The ways to minimize the computation should be a major criteria, because of the limited processor available for our system. The following table shows the analysis of the existing techniques in Vehicle health monitoring.

In order to arrive at the best strategy for this problem, it is necessary to understand the techniques used to arrive at the solution. The on board sensor collects all the data and performs partial data mining. The transfer of entire data to the site would occupy more bandwidth. It is designed to do principle component analysis, incremental Fourier transform and online linear segmentation

A. *VEDAS*

The vehicle has number of sensors. To summarize the technology used in VEDAS. Each sensor in the vehicle produce a data stream that is converted into a covariance matrix. Tridiagonalization is done through Householder method and QL method to predict the eigenvectors. After this process the clustering is performed to analyse the health pattern of vehicle monitoring. Incremental Kmeans algorithm is used for this purpose.

The control base station does the visualization of data, event based control of the vehicle. VEDAS [1] suggests the use of PCA projected data and incremental cluster, create cluster using computational geometry.

1) ***Data Projection:*** First Principle Component Analysis is done over the covariance matrix to reduce the dimensionality. First covariance matrix converted into triangular form using Householder transformation and Eigen analysis performed using QL algorithm with implicit shift. For initial data the above procedure is carried out. Later VEDAS uses sliding window to select a sample of data from the stream of sensor data. The covariance matrix is calculated for the sample data. The uncertainty or perturbation in the covariance matrix is found using Matrix Perturbation theory [5] using Forbenius norm. The procedure of PCA is done only when there is change in Eigen vector as a change in Eigen value implies a change in spectral distribution [12]. When a matrix is found with change in spectral distribution, for the matrix Principle Component Analysis is created. [13]

2) ***Clustering:*** The paper assumes that initially when the vehicle is certified as good health, then the behaviour of vehicle can be observed and the cluster becomes stable. When the observed data point falls outside the stable cluster, then the vehicle is in bad health. For this purpose the paper suggest the use of incremental clustering. Further the on-board memory is small and it would not be possible to store the entire dataset, instead the cluster can be incremented. The cluster is created using Delaunay Triangulation based polygonization approach. When an outlier is detected in the then Gaussian distribution for the set of data samples are created. VEDAS makes use of various distribution algorithm to identify the behaviour of vehicle.

Memory and Processor capacity is very important in this technology. The process which consumes the most of computation is for computing Principle Component Analysis $O(m^3,n^3)$ and the computation complexity of Covariance matrix $O(mn^2)$. PCA is performed by using House holder transformation and QL algorithm with an implicit shift [6 and 7] . This Fast Principle Component Analysis which again has a complexity of $O(mn^2)$. Even though Principle Component Analysis is done only when there is a spectral disturbance, Covariance matrix is done during fixed interval [6]. This would be burden on the processor. When the number sensor parameter increases, the computation complexity increases.

B. MINEFLEET

[2] Minefleet also was written by the same author in 2006 by the same author. On vehicle the memory capacity and also the processor capacity available is less. To enable the efficient working of this process it is necessary to implement less power memory utilization algorithm. Correlation matrix provides the correlation between two features. The correlation matrix is calculated by normalizing the dataset. The result is a normalized matrix U. U^T.U provides the correlation matrix.

In real life not all the features are correlated, at most two or three variable are correlated with one another. So it is not necessary to compute the entire correlation matrix. Instead this research paper adopts divide and conquer strategy [8]. The subsets of data are checked to find whether they have significant coefficient. Only for the subset of data correlation coefficient is found in the large sparse matrix.

When compared VEDAS, MineFleet is more suitable for continuous low overhead monitoring.

C. Vehicle Failure Prediction Using Warranty and Telematics Data [3]

In 2011 Last, Mark analyses Vehicle Failure Prediction. Previous collected crash data from SAWUR [9 and 10] (Situation-Awareness With Ubiquitous data mining for Road safety) database is used in this paper to arrive at the prediction model to minimize the risk of unexpected failure. Further when car undergoes maintenance it's periodic and based on the mileage of car. Vehicle condition do not necessarily deteriorate with time and run. This paper also aim at providing conditional maintenance, i.e. maintenance when vehicle health is not good.

Multi target Info-Fuzzy Network [11] and Weibull reliability analysis is used to predict the crash rate of vehicle. Weibull reliability analysis is based on car age, months in service and mileage. The analysis results in model of rules estimating the probability distribution of target attributes.

This technique address the crash of vehicle caused by vehicle health deterioration. It does not in scenario such as drunken drive or driver sleeps off in the middle of a long drive. It is far effective when the computation and memory is considered, but does not address all the crash scenarios.

D. On-board vehicle data mining, social networking, advertisement [4]

In 2013, the same author Kargupta published a patent, US 8478514 B2. This used the same technology as MineFleet but improved it with using Social networking. Same as MineFleet the On-board module is equipped with embedded devices, PDA to perform data stream mining. Analytics are generated from the vehicle are sent in predefined intervals to server over wireless network. The extensive analysis of data would be done in control base and further on the web server. The web server would provide a targeted advertise of the performance graph to the vehicles.

This paper improves the vehicle health monitoring system by performing mining at server and advertising the analysis to the vehicles.

IV. CONCLUSIONS

This survey paper aims at analysis of the research paper involved in "Vehicle Health Monitoring and Prediction". Paper [1], [2], [3] and [4] where studied. [1] was among the initial papers to discuss the concept of Vehicle Health Monitoring. All the possible scenarios of vehicle crash are studied in this paper. The cost of creating a covariance matrix and Principle Component Analysis where high. The author Kargupta, provided a more efficient approach in [2]. [3] discuss another approach where the data mining is distributed. The limitation of this model is that it cannot predict all cases of vehicle crash. Kargupta proposed another model for Vehicle Health monitoring in [4]. The monitoring and prediction system worked same as [2], in addition this model also sent sensor data to control server and web server for extensive data mining. The analytics found were advertised to the corresponding vehicles.

From the analysis done, [4] was found to be more suitable model for vehicle health monitoring and prediction. The Fast Principle Component Analysis discussed in [14], promises better and faster method for finding Principle Component Analysis. Model discussed in [1], [2] and [4] requires the creation of covariance or correlation matrix. Creation of covariance matrix takes time and processor capacity. [14] provides a method to directly arrive at Principle Component Analysis. According to the test performed by the author of [14], Fast Principle Component Analysis is faster than Householder transformation with QL Implicit shift, Adaptive Principle Component Analysis and Lancoz method.

REFERENCES

[1] Kargupta, Hillol, et al. "VEDAS: A mobile and distributed data stream mining system for real-time vehicle monitoring." Proceedings of the 2004 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, 2004.

[2] Kargupta, Hillol, et al. "On-board vehicle data stream monitoring using minefleet and fast resource constrained monitoring of correlation matrices." New Generation Computing 25.1 (2006): 5-32.

[3] Last, Mark. "Vehicle failure prediction using warranty and telematics data." Learn. Vol. 29. No. 3. 2011.

[4] Kargupta, Hillol. "Onboard vehicle data mining, social networking, advertisement." U.S. Patent No. 8,478,514. 2 Jul. 2013.

[5] Stewart, Gilbert W. "Matrix perturbation theory." (1990).

[6] Basilevsky. Statistical factor analysis and related methods: theory and applications. Wiley, New York, 1994

[7] Press W. [et al.]. Numerical recipes in C : the art of scientific computing. Cambridge University Press, Cambridge, 1988

[8] Alon, N., Babai, L. and Itai. A., "A Fast and Simple Randomized Parallel Algorithm for the Maximal Independent Set Problem", Journ. Of Algorithms, 7, 4, pp. 567–583, 1986.

[9] Alqallaf, F., Konis, K., Martin, R. and Zamar, R., "Scalable Robust Covariance and Correlation Estimates for Data Mining," in Proc. of the eighth ACMSIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 14–23, ACM Press, 2002.

[10] Krishnaswamy, Shonali, et al. "Towards situation-awareness and ubiquitous data mining for road safety: Rationale and architecture for a compelling application." (2005).

[11] Last, Mark. "Multi-objective classification with info-fuzzy networks." European Conference on Machine Learning. Springer Berlin Heidelberg, 2004.

[12] Patterson N, Price AL, Reich D (2006) Population Structure and Eigenanalysis. PLoS Genet 2: e190.

[13] Jolliffe, Ian. Principal component analysis. John Wiley & Sons, Ltd, 2002.

[14] Storer, Markus, et al. "Fast-robust PCA." Image Analysis (2009): 430-439.