Multi-sensor Fusion Method Using Bayesian Network for Precise Vehicle Performance Evaluation

Guillermo Alfonso De la Torre-Gea and
Gabriela García-Manzo
Institute for Research and Technology Development GARMAN A.C., Av. Andamaxei 64-40, Corregidora, Qro., Mexico. C.P. 76910.
gtorre@abanet.mx

Abstract— **In last five years, recent studies show how communication capabilities should be supported by artificial intelligence systems capable of automating many of the decisions to be taken by performance services. In this study, we propose an estimation method performance driving based on of fuel consumption and its relationship with motor sensors from vehicle information through OBD-II. We used K2 algorithm to find a configuration that best described the performance in 2.0, 2.4, and V6 3.2 motors observed data. we obtained that 2.0 and 2.4 motors present major performance than V6 motor only in short distance and low velocity. Due all variables are related developing a complex non-linear and dynamic system, Bayesian Network models are an ideal method for used in studies to require multivariable analysis. The BN models can be use as analyze the driving performance.**

Keywords—Vehicle Sensors, Driving performance, Machine Learning

I. INTRODUCTION

New communication technologies integrated into modern vehicles offer an opportunity for better assistance to people. Recent studies show how communication capabilities should be supported by artificial intelligence systems capable of automating many of the decisions to be taken by performance services.

Smali *et al.* [1] present a multi-sensor fusion strategy for a novel road-matching method designed to support real-time navigational features within advanced driving-assistance systems. The multi-sensor fusion and multi-modal estimation was realized using Dynamical Bayesian Network. Experimental results, using data from Antilock Braking System (ABS) sensors, a differential Global Positioning System (GPS) receiver and an accurate digital roadmap, illustrated the performances of this approach, especially in ambiguous situations.

Kranz *et al.* [2] present a concept for an open vehicular data interface and describe its components and architecture. They discuss the enabled applications in the context of advanced driver assistance systems with a focus on human machine interfaces, vehicle-to-x (V2X) communication and context inference systems.

Lee *et al.* [3] proposed the estimation method of fuel consumption from vehicle information through OBD-II. They assumed RPM, TPS had a relationship with fuel consumption. They got the output as fuel-consumption from a vehicle RPM, TPS as input by using polynomial equations. They had modeling as quadric function and surface function with OBD-II data and fuel consumption data supported by automotive company in real. In order to verify the effectiveness of proposed method, 5 km road-test was performed. The results showed that the proposed method can estimate precisely the fuel consumption from vehicle multi-data. It was observed that the proposed models using instantaneous engine RPM, TPS and (RPM, TPS) can predict the fuel consumption quite well with the coefficient of determination were 76%, 88% and 71% respectively.

Fogue *et al.* [4] proposes a novel intelligent system which is able to automatically detect road accidents, notify them through vehicular networks, and estimate their severity based on the concept of data mining and knowledge inference. Their system considers the most relevant variables that can characterize the severity of the accidents (variables such as the vehicle speed, the type of vehicles involved, the impact speed, and the status of the airbag). Results show that a complete Knowledge Discovery in Databases (KDD) process, with an adequate selection of relevant features, allows generating estimation models able to predict the severity of new accidents.

Meiring [5] evaluates the possibilities for unique driver identification utilizing the approaches identified in other driver behavior studies. It was found that Fuzzy Logic inference systems, Hidden Markov Models and Support Vector Machines consist of promising capabilities to address unique driver identification algorithms if model complexity can be reduced.

Road safety problems are a major area of concern in the transport industry, especially in Developed countries. Traffic accident involvement is more closely related to human judgement and decision-making than the mere inability to control the vehicle, and therefore, the focus of driver behavior and decision-making patterns became a popular research area in road-safety applications [6].

In [7], driving style classification between four driving styles (normal, drunk, reckless, and fatigue) is performed in real time by introducing a probabilistic model based on Dynamic Bayesian networks (DBNs). Contextual information regarding the driver, the vehicle, and the environment are fused together. The dynamic behavior model captures the static and the temporal aspects of the driver's behavior and led to robust and accurate behavior detection.

The k-Nearest Neighbours (k-NN) algorithm is a well-known nonparametric classifier which is based on simple and effective supervised learning and classification techniques. Supervised locally linear embedding (SLLE) proved to be a powerful feature extraction method in the feature extraction for plant leaf recognition (8). A recognition rate of greater than 90% was achieved on five kind of plants using an input vector for the k-NN algorithm. The extraction of driving event features are necessary and the SLLE method should be investigated.

Enev *et al.* [9], investigate the potential to identify individuals using sensor data snippets of their natural driving behavior. They split the data into training and testing sets, train an ensemble of classifiers, and evaluate identification accuracy of test data queries. Their results indicate that, at least among small sets, drivers are indeed distinguishable using only in car sensors.

Thus, the aim of this study was to evaluate the sensors relationship of three different motors in real time, using Bayesian analysis of a dynamic data; to determine the best performance to design a solution of vehicular park.

II. THEORY

New communication technologies integrated into modern vehicles offer an opportunity for better assistance to people. Recent studies show how communication capabilities should be supported by artificial intelligence systems capable of automating many of the decisions to be taken by performance services.

Smali *et al.* [1] present a multi-sensor fusion strategy for a novel road-matching method designed to support real-time navigational features within advanced driving-assistance systems. The multi-sensor fusion and multi-modal estimation was realized using Dynamical Bayesian Network. Experimental results, using data from Antilock Braking System (ABS) sensors, a differential Global Positioning System (GPS) receiver and an accurate digital roadmap, illustrated the performances of this approach, especially in ambiguous situations.

Kranz *et al.* [2] present a concept for an open vehicular data interface and describe its components and architecture. They discuss the enabled applications in the context of advanced driver assistance systems with a focus on human machine interfaces, vehicle-to-x (V2X) communication and context inference systems.

Lee *et al.* [3] proposed the estimation method of fuel consumption from vehicle information through OBD-II. They assumed RPM, TPS had a relationship with fuel consumption. They got the output as fuel-consumption from a vehicle RPM, TPS as input by using polynomial equations. They had modeling as quadric function and surface function with OBD-II data and fuel consumption data supported by automotive company in real. In order to verify the effectiveness of proposed method, 5 km road-test was performed. The results showed that the proposed method can estimate precisely the fuel consumption from vehicle multi-data. It was observed that the proposed models using instantaneous engine RPM, TPS and (RPM, TPS) can predict the fuel consumption quite well with the coefficient of determination were 76%, 88% and 71% respectively.

Fogue *et al.* [4] proposes a novel intelligent system which is able to automatically detect road accidents, notify them through vehicular networks, and estimate their severity based on the concept of data mining and knowledge inference. Their system considers the most relevant variables that can characterize the severity of the accidents (variables such as the vehicle speed, the type of vehicles involved, the impact speed, and the status of the airbag). Results show that a complete Knowledge Discovery in Databases (KDD) process, with an adequate selection of relevant features, allows generating estimation models able to predict the severity of new accidents.

Meiring [5] evaluates the possibilities for unique driver identification utilizing the approaches identified in other driver behavior studies. It was found that Fuzzy Logic inference systems, Hidden Markov Models and Support Vector Machines consist of promising capabilities to address unique driver identification algorithms if model complexity can be reduced.

Road safety problems are a major area of concern in the transport industry, especially in Developing countries. Traffic accident involvement is more closely related to human judgment and decision-making than the mere inability to control the vehicle, and therefore, the focus of driver behavior and decision-making patterns became a popular research area in road-safety applications [6].

In [7], driving style classification between four driving styles (normal, drunk, reckless, and fatigue) is performed in real time by introducing a probabilistic model based on Dynamic Bayesian networks (DBNs). Contextual information regarding the driver, the vehicle, and the environment are fused together. The dynamic behavior model captures the static and the temporal aspects of the driver's behavior and led to robust and accurate behavior detection.

The k-Nearest Neighbors (k-NN) algorithm is a well-known nonparametric classifier which is based on simple and effective supervised learning and classification techniques. Supervised locally linear embedding (SLLE) proved to be a powerful feature extraction method in the feature extraction for plant leaf recognition (8). A recognition rate of greater than 90% was achieved on five kinds of plants using an input vector for the k-NN algorithm. The extraction of driving event features are necessary and the SLLE method should be investigated.

Enev *et al.* [9], investigate the potential to identify individuals using sensor data snippets of their natural driving behavior. They split the data into training and testing sets, train an ensemble of classifiers, and evaluate identification accuracy of test data queries. Their results indicate that, at least among small sets, drivers are indeed distinguishable using only in car sensors.

Thus, the aim of this study was to evaluate the sensors relationship of three different motors in real time, using Bayesian analysis of a dynamic data; to determine the best performance to design a solution of vehicular park.

III. Materials and Methods

We obtained four datasets with 2644 registers of 60 s to monitored route (Fig. 1) over an OBD2 connector, in which four cars with different motor characteristics: 2.0 automatic, 2.4 automatic, 2.4 standard, and 3.2 automatic were conducted, with the aim of determine the values of their sensors.

To analyzing the relationship between motor sensors, we used ELVIRA software v 0.162 in three stages, suggested by Ortiz-Vazquez [15]:

1) It is carried out using the algorithm of allocation "to mean" to complete the series of partial data. This algorithm replaces lost or unknown values, the mean values for each variable. This method requires no limits and involves discretizing the massive data by the algorithm using six intervals with the same frequency.

2) According to Espinoza-Huerta *et al.* [16], the best Bayesian network structure is developed using K2 algorithm with 5 parents and without restrictions.

3) We performed dependency analysis to get the topological structure of the network, which represents the causal variables and their dependencies. After obtaining a parametric learning network, we calculated the conditional probabilities variables that show the relationship or dependence.

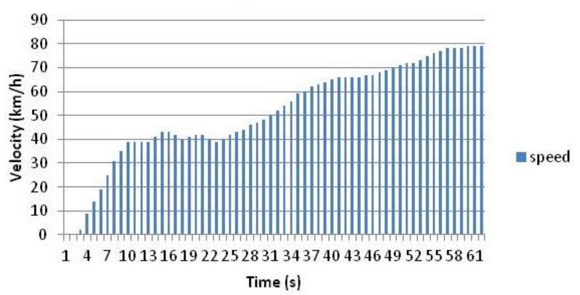

Fig. 1. *Road-test evaluation.*

IV. Results and Discution

In Fig 2, we present a BN model to indicate the behavior of automatic turbo charged motor sensors, in which each sensor is a probabilistic variable, and each relationship is indicate with an arc; red arc indicates a directed relationship and blue arc is an inverse relationship.

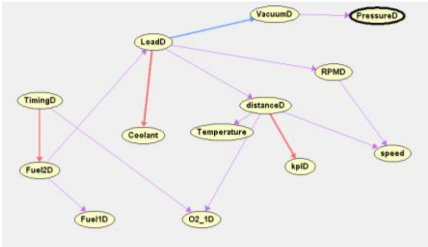

Fig. 2. Relationship between all variables over an automatic turbo charged 4 cylinders 2.4 motor.

Fuel consumption (Fuel2), coolant temperature and Kilometers per liter (KPL), are directly dependent of the Timing sensor, Load and Distance respectively; and Vacuum is inversely dependent to the Load. Other variables were not indicated a defined and measured relationship, but all variables are related developing a complex non-linear and dynamic system. This model only applies in specific conditions showed in Fig 1.

In Figure 3, we noted the difference with the same car was conducted in standard mode; in which the distance is inversely dependent of the load and directly influences the vehicle speed. In other hand, coolant temperature is inversely dependent to the oxygen sensor; which means that in cool motor conditions, O_2 sensor sends a “close loop” signal.

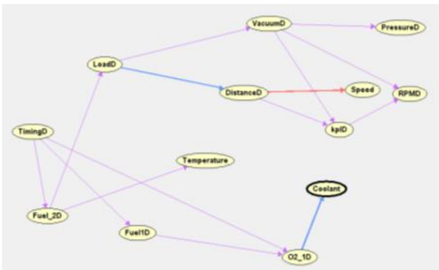

Fig. 3. *Relationship between all variables over a standard turbo charged 4 cylinders 2.4 motor.*

When we compared 2.4 with 2.0 motor, we noted that distance was directly dependent to load, as show in Fig. 4 that means second car presents major performance.

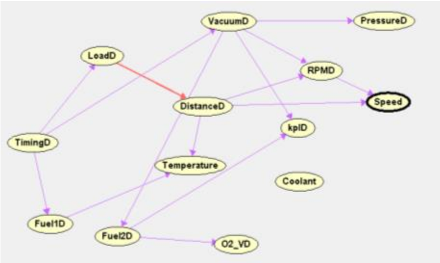

Fig. 4. *Relationship between all variables to an atmospheric 4 cylinders 2.0 motor.*

When we conducted the 2.4 turbo charger motor in manual mode, the fuel consumption is high because synchronization is not the best. Finally, human errors are present when we increase our participations in processes with computing control.

In Fig. 5, BN model to V6 motor shows relationship which fuel consumption is inversely dependent to distance, kpl, and speed. In the other hand, timing inversely influences both: Load and RPM. This model indicates that V6 motor has less performance in short distance and low velocity due requires major fuel consumption. But those relationships indicate too that in major distances, V6 has high performance.

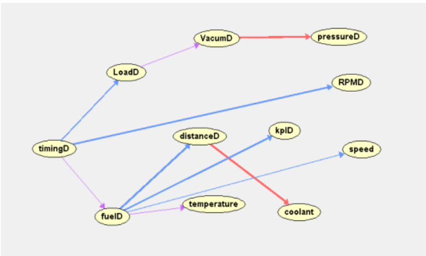

Fig. 5. *Relationship between all variables over a v6 atmospheric 3.2 motor.*

Those results are only applies to specific conditions showed in Fig 1.

V. Conclusion

Performance driving depends to driving modes at specific conditions, and then its measure is limited. For those study conditions, we obtained that 2.0 and 2.4 motors present major performance than V6. Manual mode driving consumption is high because synchronization is not the best. V6 motor has less performance in short distance and low velocity. Due all variables are related developing a complex non-linear and dynamic system, Bayesian Network models are an ideal method for used in studies to require multivariable analysis. The BN models can be use as analyze the driving performance.

Acknowledgment

The Mexican National Council of Science (CONACYT) supported this research.

References

[1] Smaili1, C., El Najjar M., Charpillet, F. 2007. Multi-Sensor Fusion Method using Dynamic Bayesian Network for Precise Vehicle Localization and Road Matching. DOI: 10.1109/ICTAI.2007.70 · Source: IEEE Xplore.

[2] Kranz, M., Weber, E., Frank, K., Hermosilla, D. 2009. Open Vehicular Data Interfaces for In-Car Context Inference. Proceedings of the First International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive UI 2009), Sep 21-22 2009, Essen, Germany.

[3] Lee, M.G., Park, Y.K., Jung, K.K., Yoo, J.J. 2011. Estimation of Fuel Consumption using In-Vehicle Parameters. International Journal of u- and e- Service, Science and Technology Vol. 4, No. 4,: 31 – 46.

[4] Fogue, M., Garrido, P., Martínez, FJ., Cano Escribá, JC., Tavares De Araujo Cesariny Calafate, CM., Manzoni, P. (2014). A system for automatic notification and severity estimation of automotive accidents. IEEE Transactions on Mobile Computing. 13(5):948-963. doi:10.1109/TMC.2013.35.

[5] Meiring, G.A.M., Myburgh, C.H. 2015. A Review of Intelligent Driving Style Analysis Systems and Related Artificial Intelligence Algorithms Sensors, 15: 30653–30682.

[6] French, D.J., West, R.J., Elander, J., Wilding, J.M. Decision-making style, driving style, and self-reported involvement in road traffic accidents. Ergonomics **1993**, 36, 627–644.

[7] Al-Sultan, S., Al-Bayatti, A., Zedan, H. Context-aware driver behavior detection system in intelligent transportation systems. IEEE Trans. Veh. Technol. **2013**, 62, 4264–4275.

[8] Zhang, S., Chau, K.W. Dimension Reduction using Semi-supervised Locally Linear Embedding for Plant Leaf Classification. In Proceedings of the 5th International Conference on Intelligent Computing, Ulsan, Korea, 16–19 September 2009; pp. 948–955.

[9] Enev, M., Takakuwa, A., Koscher, K., Kohno, T. Automobile Driver FingerprintingProceedings on Privacy Enhancing Technologies ; 2016 (1):34–51

[10] Hruschka E, Hruschka E, Ebecken N F F, *Bayesian networks for imputation in classification Problems*, J Intell Inform Syst, 29 (2007) 231.

[11] Reyes P, Bayesian networks for setting genetic algorithm parameters used in problems of geometric constraint satisfaction, Intel. Artificial, 45 (2010) 5.

[12] Garrote L, Molina M, and Mediero L. Probabilistic Forecasts Using Bayesian Networks Calibrated with Deterministic Rainfall- Runoff Models, in Vasiliev *et al*. edited by Extreme Hydrological Events: New Concepts for Security (Springer) 2007, 173.

[13] Borunda, M., Jaramillo, O. A., Reyes, A., Ibargüengoytia, P.H. (2016). Bayesian Networks in Renewable Energy Systems: A Bibliographical Survey. Renewable and Sustainable Energy Reviews, 62; 32–45.

[14] Guoiang, L. Knowledge Discovery with Bayesian Networks, Ph. D. thesis, National University of Singapore, Singapore, 2009.

[15] Guillermo de la Torre-Gea, Oscar Delfín-Santisteban, Irineo Torres-Pacheco, Genaro Soto-Zarazúa, Ramón Guevara-González, Enrique Rico-García. *Bayesian networks applied in a CFD model of the crop in greenhouse.* Agrociencia, 48(2): 307 – 319, (2014).

[16] Ortiz-Vazquez, I.C., Pérez-Robles, J.F., Fernandez-Loyola, R., Pérez-Brito, J.F., De La Torre-Gea, G.A. *A Multivariable Computational Fluid Dynamics Analisys Method Based In Bayesian Networks Applied In A Bioreactor.* Journal of Applied Chemical Science International, 6(1): 10-17, (2016).

[17] Espinoza-Huerta, T.D., Ortíz-Vázquez, I.C., García-Manzo, G., De la torre-Gea, G.A. *A Multivariable Computational Fluid Dynamics Validation Method Based in Bayesian Networks Applied in a Greenhouse.* International Journal of Agriculture Innovations and Research. (4),1: 67 – 71, (2015).