Fault Diagnostics Using Network Motif Signature

Tsai-Ching Lu[1], David L. Allen[1], Yilu Zhang[2], and Mutasim A. Salman[2]

[1]HRL Laboratories LLC, Malibu, California, 90265, U.S.A

tlu@hrl.com
dlallen@hrl.com

[2]General Motors, Warren, Michigan, 48090, U.S.A

yilu.zhang@gm.com
mutasim.a.salman@gm.com

ABSTRACT

In modern vehicles, controls are distributed over multiple Electronic Control Units (ECUs) that are connected through in-vehicle communication networks. Fault diagnostics for such a distributed control system is very challenging, which has resulted in many no-trouble-found (NTF) cases during warranty repairs. To address this problem, we propose a novel network-theoretic approach that detects, identifies, and localizes faults using both the structure of the communication network (topological information) and message flow information. The proposed method not only enables the characterization of normal operation and a-priori known faults across communication networks, which is already beyond the current practice of individual ECU centric diagnostics, but also the diagnostics of unknown or cascading failures emerging from unexpected operational environments.

1. INTRODUCTION

In-vehicle electrical and electronic systems (EES) are embedded systems that implement advanced fuel-economy, emission control, safety, convenience, and comfort features. Although EES systems are largely deterministic with designed subsystem and component interactions to accomplish desired functionality, it may become stochastic when vehicles operate in extreme and unexpected conditions with fault propagations that were not anticipated during the design, testing, or validation stages. It is the emerging faults and cascading failures that pose ever growing challenges to diagnostics and prognostics in complex electronics, especially in complex systems such as in-vehicle EES.

The state-of-the-art diagnostic techniques for in-vehicle ESS are generally based on Diagnostic Trouble Codes (DTC). Each individual ECU records the DTCs when internal, local, error-checking routines detect fault conditions. Technicians then inspect DTC archives, and perform diagnostic tests to determine the root causes. In practice, one fault may be detected by multiple ECUs and trigger multiple DTCs. One single fault may also cascade from one part of the system to other parts of the system, and trigger multiple DTCs. In either case, it has been a significant challenge to identify the root cause of multiple observed DTCs. This situation is particularly challenging for communication related DTCs. Moreover, the error checking routines in individual ECUs usually consider only faults anticipated during the design phase, which don't include unknown failures or cascading failures that emerge from unexpected operational conditions.

In this paper, we propose a novel network-theory approach to identify and localize faults based on topological information of the in-vehicle communication network, network motif fault signatures, and message flow information over the network. The proposed methodology is an enabler to

1. Characterize normal operation and a-priori known faults across communication networks which is beyond the current myopic view of observability employed by individual ECUs;
2. Identify and localize unknown or cascading failures emerging from unexpected operational conditions.

The rest of the paper is organized as follows. After a brief survey of related research in Section 2, we present the proposed approach in Section 3, and follow up with technical description for each module of the approach in Sections 4-7. We then report our initial empirical study of the proposed methodology on a multi-ECU system simulated in Vector CANoe (vector.com) in Section 8.

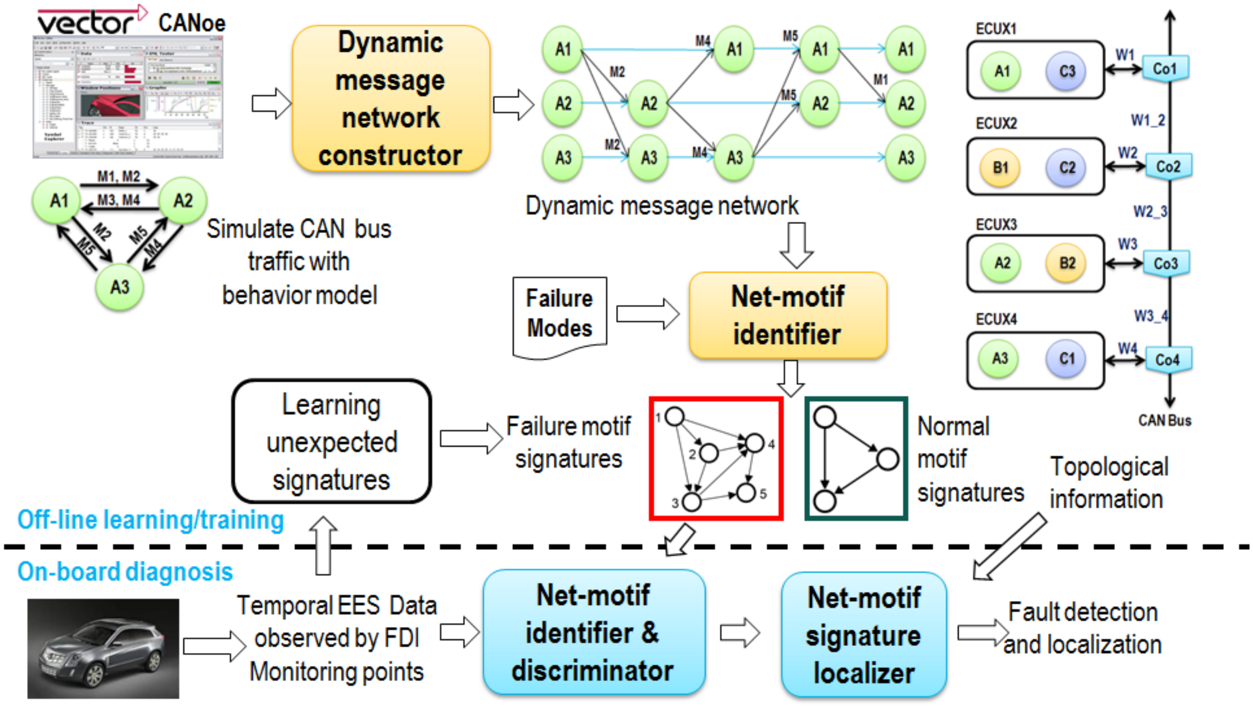

Figure 1. Normal motif and fault motif signatures are trained during the off-line learning/training stage to create a detector and localizer that can be deployed into diagnostic software tools or real-time diagnosis assistance on vehicles. On-board fault detection and isolation (FDI) monitoring point is selected based on methodology in Lu et al. (2011).

2. RELATED RESEARCH

Net-motifs are sub-graphs of fixed size or length which appear more commonly than others in a network. In this paper, net-motifs represent the observation of message transmission patterns on the communication bus.

Milo et al. (2002) first introduced the concept of network motifs and demonstrated their abundance in biochemistry (transcriptional gene regulation), ecology (food webs), neurobiology (neuron connectivity), and engineering (electronic circuits, World Wide Web). The net-motif concept has not been applied to in-vehicle communication networks of EES. Rather than focusing on the existences of network motifs in engineering field, we are taking the net-motifs, combining with topological information, to perform fault identification and localization.

Wang et al. (2009) described a network management system to diagnose network traffic faults for TCP/IP communication networks. Their model is trained to derive fault signatures from temporal patterns in historical network traffic data, in conjunction with network topological information. Topological information is demonstrated for selfing/neighboring, containing/contained, down/up-streaming, and tunneling. The key concept is utilizing the event bursts for steps of fault signature learning and indexing. The term motif is mentioned, but not specifically utilized. Only pair-wise event interactions are considered. There is no mentioning of in-vehicle networks and behavior models.

Dijev et al. (2011) introduced graph based statistical analysis on network traffic of internet for intrusion activity or malicious behaviors. The telescoping graph is proposed to capture the decomposition of a protocol graph which describes activity between hosts in the network for a given protocol, e.g. SSH protocol graph constructed from packet header. A discrete hazard model is learned to detect anomalies from decomposed TSG graphs. Although statistical graph-based methods are used, there is no direct application to in-vehicle networks, and there is no consideration of topological information.

3. SYSTEM AND METHODOLOGY

We develop a system and methodology that assists with fault identification and localization for in-vehicle embedded electrical and electronic systems (EES) using network motif signatures. The system takes behavior models, topological information, and communication traffic as input, and outputs net-motif-based fault signatures and the candidate fault sets, if emerging or a-priori known failures occurs in vehicles. The system is first trained off-board to characterize net-motif signatures under normal operations and known failure-modes in design phase, and later deployed on-board or embedded in a troubleshooting tool to assist failure net-motif signature detection and fault localization (Figure 1). The details of each module in Figure 1 are described in the following sections: Dynamic Message Network Constructor (Section 4), Net-motif Identifier (Section 5), Net-motif Signature Detector (Section 6), and Net-motif Signature Localizer (Section 7). Figure 2 shows the data flow of the proposed fault detection and localization using net-motif signatures.

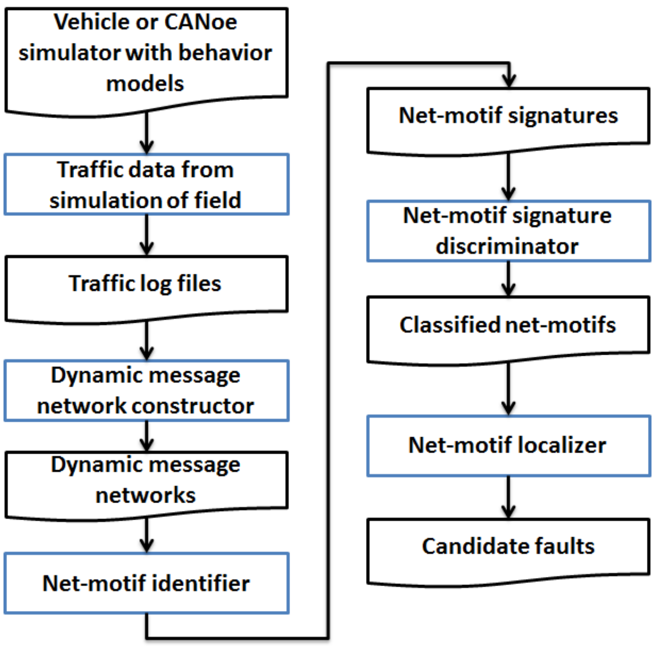

Figure 2. *Dynamic message network constructor* takes traffic log files to generate dynamic message networks; *Net-motif identifier* finds network motifs in constructed dynamic message networks; *network-motif signature detector* compares identified motif profiles against normal operation motif profiles to classify fault motifs; *net-motif signature localizer* employs topological information and behavior models (designed message flows) to generate candidate fault sets.

4. Dynamic Message Network Construction

We develop the dynamic message network constructor to take network traffic as input, and output a dynamic message network where nodes represent ECUs and edges represent messages flows between ECUs. The pseudo code for one instantiation of constructing a dynamic message network is described as follows.

1. Initialize discrete counter T=1;
2. For each message [ECU_i -> $ECU_{j,\,k,\,\ldots}$]
   a. Let t_{tx} = T;
   b. If the sender of the current message is found as a receive-only node (with only incoming edges) within the previous W seconds, counted from the current message timestamp (simulation or real time values)
      i. Set t_{tx} equal to the counter value associated with ECU_i; change ECU_i to transmit node from receive-only node;
   c. Else
      i. create a new ECU_i transmit node and associated it with the counter value t_{tx}
   d. Create receive nodes $ECU_{j,\,k,\ldots}$ if they do not already exist at the counter value t_{tx}+1 (note this may not be T+1 if used previous receive node);
   e. Add edges from associated ECU_i -> $ECU_{j,\,k,\,\ldots}$ node
   f. Set T = t_{tx}+1 (the receive node's counter value of the current message)

An example dynamic message network generation is shown in Figure 3.

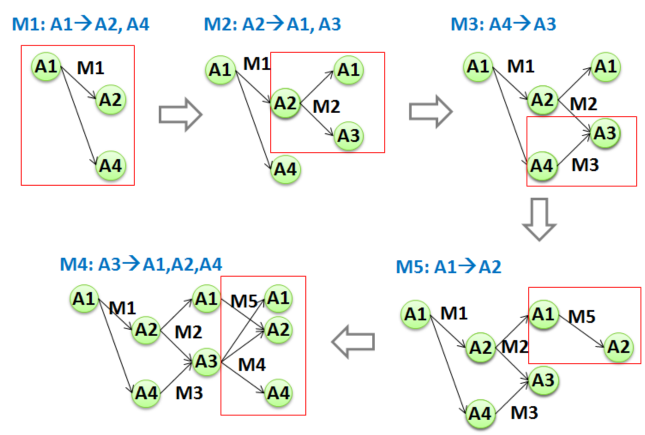

Figure 3. Dynamic message network generation: Messages M1, M2, M3, M5 and M4 are detected and corresponding networks are generated after processing each message. Due to the sliding window parameter, the message M3 from A4 to A3 can have exchangeable temporal sequence with message M2 from A2 to A1 and A3, i.e., two directed edges into A3 before A3 can respond to send out M4.

The dynamic message network constructor will be used first in off-line processing to establish message networks for identifying normal operational net-motifs and net-motif fault signatures. The constructor will later be utilized in on-board diagnosis. The parameter W second is a tunable such that dynamic message networks could be adapted to different communication architectures in vehicles.

5. Net-Motif Signature Identification

Taking a dynamic message network as input, the net-motif identifier detects net-motifs by extracting and grouping sub-graph patterns in the network. The output of the net-motif identifier are detected net-motifs with counts providing a distribution of motifs. We build our net-motif identifier based on the efficient motif discovery algorithm in (Wenick, 2006) and efficient sub-graph isomorphism algorithms in (Cordela et al., 2004).

In the extraction step, we will enumerate size-k sub-graphs by growing a recursion tree (RT) as follows:

1. Assign nodes of input graphs with ordered indices
2. Each node in a RT will be associated with two sets sub-graph sets (Vsub) and exclusive neighbors (Vecn) except the root node.
3. Nodes in Vsub are called spawning nodes and nodes in Vecn are nodes with indices greater than the associated spawning nodes in Vsub.
4. Recursively grow the RT to the k-level which would be the sub graph with size k.

In the grouping stage, we implement graph isomorphism algorithms in (Cordela et al., 2004) to group extracted sub-graphs into canonical patterns of net-motifs. Figure 4 (a) and (b) illustrate the steps in net-motif extraction with recursive tree and identified motifs for an example message network.

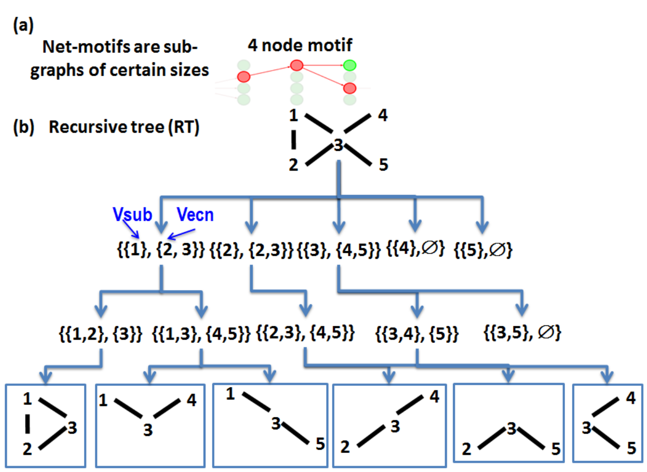

Figure 4. Illustration of Net-motif identification. (a) shows the example 4-node net-motifs in a fragment of dynamic message network; (b) shows the recursive tree constructed to identify 3-node net-motifs in the example 5 node network;

6. Fault Detection Using Net-Motif Signature Discrimination

Given net-motif distributions, we will learn a net-motif discriminator to classify motifs into different modes of operations, such as normal and failures, and detect net-motif of failure operations by comparing against motifs of normal operations to produce fault motif signatures.

We train net-motif discriminator as follows:

1. Simulate or collect network traffics under normal operational conditions
2. Inject failures in network traffic simulation tools, such as CANoe, or inject failures onto real vehicles to collect fault traffic for known failure modes
3. Perform dynamic message network construction to build network for collected traffic data
4. Perform net-motif identification to derive net-motif distributions from normal and failure dynamic message networks.
5. Learn net-motif discriminator on net-motif distributions using machine learning algorithms, such as probabilistic graphical models or other classification algorithm such as k-nearest neighbor or Support Vector Machines.
6. Perform cross-validation to validate the performance of trained net-motif discriminator.

The output of net-motif discriminator is categories of net-motif signatures under normal or failure conditions.

7. Fault Localization Using Net-Motif Signatures

Although the trained net-motif discriminator could achieve high detection and classification accuracy for known failure modes, it is impossible to simulate all combinations of all known failure modes in various failing sequences. Moreover, unknown failures may occur when vehicle operating in extreme or unexpected conditions. In this section, we describe the net-motif failure signature (NMFS) localizer that detects and localizes unknown faults via net-motifs and topological information.

The topological information is provided as an input to NMFS in the form of expected message flows from behavior models. Each message is associated with a sequence of flow paths starting from a sender ECU, going through sequences of wires, connectors, and gateways to finally reach the receiver ECU. For example, ECU1 sends message M1 on the bus and ECU2 and ECU3 are in the same virtual networks which are supposed to listen and receive message M1. We will write topological information of message flow as {ECU1, wire1, connector1, …., wireN, ECU2} and {ECU1, wire1, connector1, …, wireI, ECU3}.

Given such topological information and profiles of normal operational motifs, we perform the following steps to do fault localization for unknown faults:

1. Perform dynamic message network construction on traffic data
2. Perform net-motif identification to extract net-motifs from the dynamic message network

3. Apply sub-graph matching to identify the deviation of monitored net-motifs with those in net-motif profiles of normal operations.
4. Trace deviations (observed or missing messages) over topological information of message flow to discern the ambiguity sets of possible faults using set intersection and differentiation operations.

The output of NMFS is an ambiguity set of likely faults for net-motif failure signatures that deviate from normal operations. Figure 5 illustrate the concept.

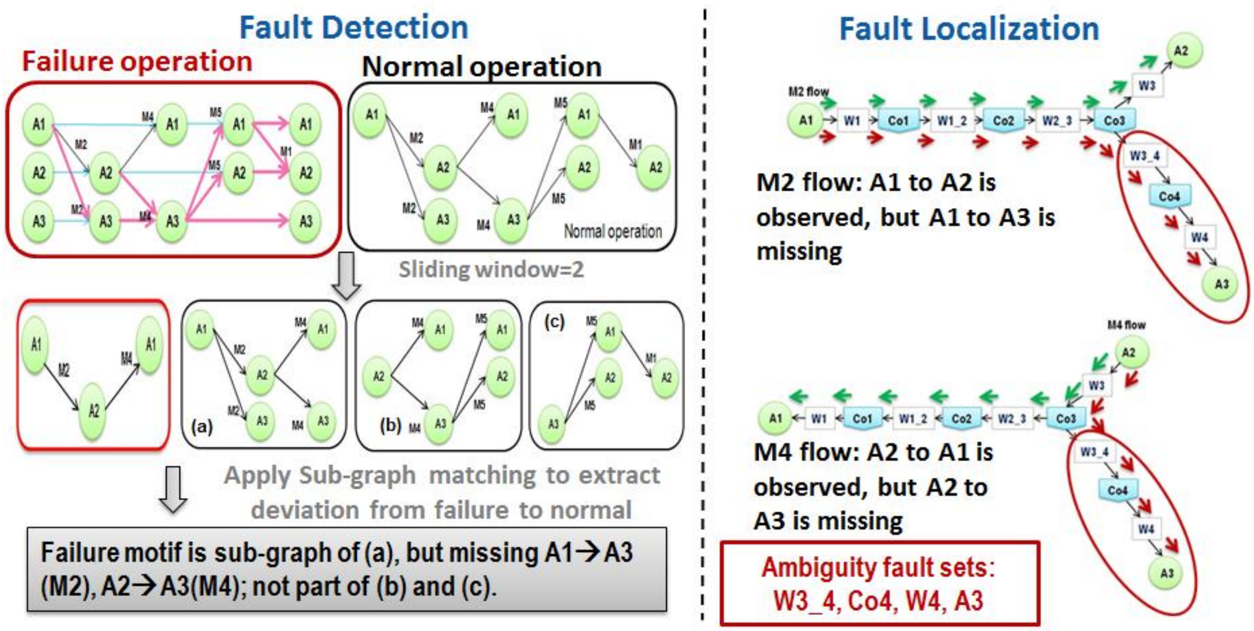

Figure 5. Fault detection localization using net-motif fault signatures. The left-panel shows the steps of fault detection which compare net-motif identified from dynamic message network of failure operation against net-motifs of normal operation. The step of fault localization applies sub-graph matching to identify the observed net-motif fault signature, which is a sub graph of a net-motif in normal operation. This leads to perform the operation of set differentiation over the message flows on topological structure (CAN bus) for observed (green) and unobserved (red) message flows, and recommend ambiguity fault sets of components such as wires , connectors, or ECUs.

8. AN EXAMPLE STUDY

We show an example communication network with 4 ECUs and two behavior features A and B. The topology is shown in Figure 6 (a) with wires (annotated with W*) and connectors (annotated with Co*). The behavior models are shown in Figure 6 (b). The canonical dynamic message networks for normal operations for the two features are shown in Figure 6 (c). We simulate 2 hardware wire failure and 4 software failure for software components in two features as in Figure 6 (d).

We simulate normal and failure operations with noise, and generate 21 runs of varying durations corresponding to 7 modes of operation with 3 runs for each mode. We perform net-motif identification over each log file and derive its net-motif distribution. We then perform net-motif discrimination to derive a similarity score by comparing pair-wise differences in the identified distributions. We use the following cosine similarity function:

$$sim(NM, FM) = \frac{\sum_i NM_i \times FM_i}{\sqrt{\sum_i (NM_i)^2}\sqrt{\sum_i (FM_i)^2}}$$

Where NM_i is the counts of one motif distribution, and FM_i is the counts of the second motif distribution. The dis/similarity of net-motifs for different modes of operations is shown in Figure 7.

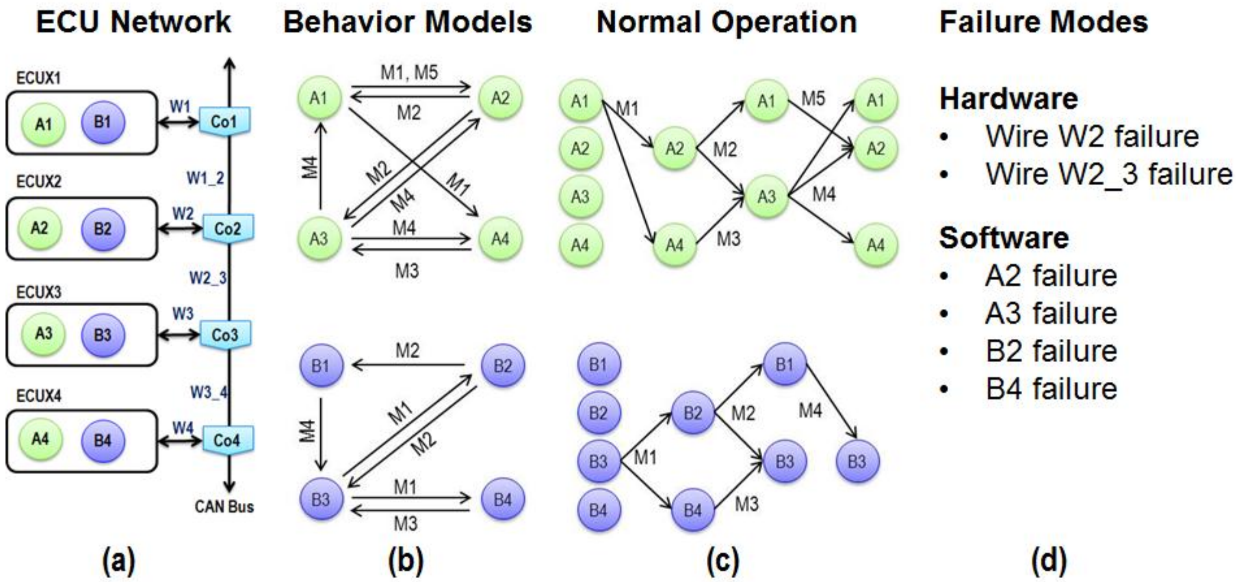

Figure 6. Simulation of communication network traffics: (a) shows 4 ECUs, wires and connectors with their topological connections; (b) shows behavior models of two features implemented on 4 ECUs; (c) shows normal operational of canonical dynamic message network for two features; (d) shows simulated hardware and software failure modes.

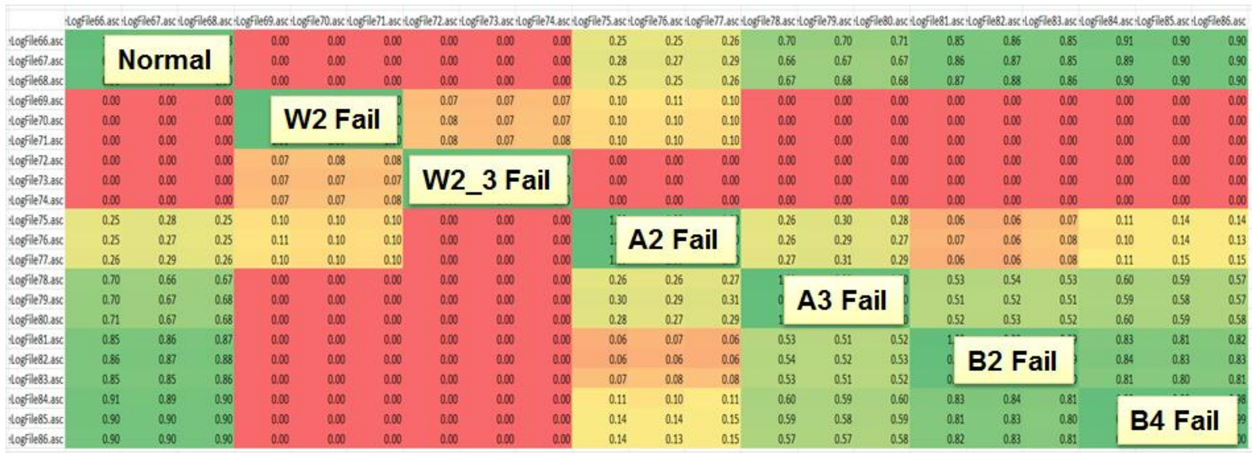

Figure 7. The resulting similarity matrix for net-motif distributions from the dynamic message constructed from simulated network traffic in Figure 6. Each row and each column represent one run of the simulation, and each mode of operation was simulated with 3 runs. The matrix is color-coded based on similarity score (green: similar, red: different). Some net-motif fault signatures (e.g. B2 Fail and B4 Fail) are fairly similar to those in normal operation, and some are quite different (e.g. W2 Fail and W2_3 fail). Despite this mixed of similarity, our net-motif discriminator based on k-nearest neighbor can discriminate different mode of operations with 100% accuracy on new log files which it was not trained on.

9. Conclusion

The increasing complexity of in-vehicle electrical and electronic systems poses ever-growing challenges to diagnostics and prognostics. In 2011, we leveraged the latest advancements in Network Science to develop a system and method based on betweenness centrality to find good monitoring points for fault detection (Lu et al. 2011). In 2012, we address the challenge of emerging or cascading faults in EES that are difficult to be addressed by traditional component-level diagnostics. We present the novel network-theoretic approach to detect, identify, and localize faults using both the structure of communication network (topological information) and message flow information over the communication network. The proposed method enables the characterization of normal operation and a-priori known faults across communication networks, which is beyond the current practice of individual ECUs point of view. It also differs from conventional signature-based approach in that the mode of operations is classified at the level of modular interactions (motifs) rather than individual observation. It also enables the diagnostics of unknown or cascading failures emerging from unexpected operational environments. The system has been demonstrated on a simulated multi-ECU system in the CANoe environment, and shows promising results in fault diagnostics.

ACKNOWLEDGEMENT

The authors would like to thank Shengbing Jiang and Mark Howell for many valuable technical discussions. Special thanks also go to many of our colleagues in Engineering, including Doug Duddles, Sandeep Menon, Ken Orlando, Bob Schwabel, Lars Soderlund, Mike Sowa, and Natalie Wienckowski, for sharing the domain knowledge and shaping the research direction.

REFERENCES

Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks, International AAAI Conference on Weblogs and Social Media.

Cordella, L.P., Pasquale, F., Sansone, and C., Vento, M. (2004). A (Sub)Graph Isomorphism Algorithm for Matching Large Graphs, , IEEE Transaction on Pattern Analysis and Machine Intelligence, 26(10), 1367-1372.

Djidjev, H., Sandine. G., and Storlie, C. (2011). Graph Based Statistical Analysis of Network Traffic, MLG.

Lu, T.-C., Zhang, Y., Allen, D. L. and Salman, S. M. (2011). Design for Fault Analysis Using Multi-Partite, Multi-Attribute Between Centrality Measures, PHM.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002), Network-motifs: Simple Building Blocks of Complex Networks, Science, vol. 298, no. 5594, pp. 824–827.

Wang, T., Srivatsa, M., Agrawal, D., and Liu, L. (2009), Learning, Indexing, and Diagnosing Network Faults, KDD.

Wernicke, S. (2006), Efficient Detection of Network Motifs, IEEE/ACM Transaction on Computational Biology and Bioinformatics, 3(4), 347-359.

BIOGRAPHIES

Tsai-Ching Lu received his M.S. degree in Computer Sciences from NYU in 1996, and his M.S. and Ph.D. degrees in Intelligent Systems from University of Pittsburgh in 2000 and 2003 respectively. He joined HRL Laboratories, LLC, in 2003 and currently holds the position of Senior Research Staff Scientist. His research interests include probabilistic causal reasoning, complex systems, and network sciences with applications to behavior modeling of socio-technological systems. He received HRL's Distinguished Inventor Award in 2010 and GM contribution Award in 2011.

Yilu Zhang received his B.S., and M.S. degrees in Electrical Engineering from Zhejiang University, China, in 1994, and 1997, respectively; and his Ph.D. degree in Computer Science from Michigan State University in 2002. He joined the R&D center of General Motors Co. at Warren, Michigan in 2002, and currently holds a position of Staff Researcher. Dr. Zhang's research interests include statistical pattern recognition, machine learning, signal processing, and their applications such as integrated system health management and human machine interactions. Dr. Zhang is a Senior Member of IEEE. He is a winner of 2008 "Boss" Kettering Award – the highest technology award in GM – for his contribution to Connected Vehicle Battery Monitor, a remote vehicle diagnostics technology.

David Allen received his M.S. and Ph.D. degrees in Computer Science from the University of California, Los Angeles (UCLA) in 2001 and 2005 respectively. He joined HRL Laboratories, LLC, in 2006 and currently holds the position of Research Staff Scientist. His research interests include artificial intelligence, probabilistic reasoning, decision making under uncertainty, complex systems analysis, and network science; he has applied them to many domains including integrated system health management, cybersecurity, social network analysis, and behavior modeling. In 2010 he received HRL's Distinguished Inventor Award and in 2011 received a GM Contribution Award.

Mutasim Salman is a Lab. Group manager and a Technical Fellow in the Electrical, Controls and Integration Lab. of GM Research and Development Center. He has the responsibility of development and validation of algorithms for state of health monitoring, diagnosis, prognosis and fault tolerant control of vehicle critical systems. He pioneered the work on integrated chassis control in the late eighties that led to the production of GM "Industry First" Stabilitrak1 and then to Stabilitrak3. He had an extensive experience in hybrid vehicle, modeling, control and energy management strategies. He has several GM awards that includes 3 GM prestigious Boss Kettering, 3 McCuen, and 2 President and Chairman Awards. Mutasim received his bachelor's degree in Electrical Engineering from University of Texas at Austin; M.S. and PhD in Electrical Engineering with specialization in Systems and control from University of Illinois at Urbana- Champaign. He also has an Executive MBA. He is IEEE senior member. He holds 21 patents and has coauthored more than 42 refereed technical publications and a book. He joined the GM R&D staff in 1984.