Frontiers of Information Technology & Electronic Engineering
www.zju.edu.cn/jzus; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

# Controller area network node reliability assessment based on observable node information*#

Lei-ming ZHANG, Long-hao TANG, Yong LEI‡

(*State Key Laboratory of Fluid Power Transmission and Mechatronics, Zhejiang University, Hangzhou 310027, China*)

E-mail: lmzhang@zju.edu.cn; tlh123@yeah.net; ylei@zju.edu.cn

Received Feb. 26, 2016; Revision accepted Sept. 14, 2016; Crosschecked Apr. 23, 2017

**Abstract:** Controller area network (CAN) based fieldbus technologies have been widely used in networked manufacturing systems. As the information channel of the system, the reliability of the network is crucial to the system throughput, product quality, and work crew safety. However, due to the inaccessibility of the nodes' internal states, direct assessment of the reliability of CAN nodes using the nodes' internal error counters is infeasible. In this paper, a novel CAN node reliability assessment method, which uses node's time to bus-off as the reliability measure, is proposed. The method estimates the transmit error counter (TEC) of any node in the network based on the network error log and the information provided by the observable nodes whose error counters are accessible. First, a node TEC estimation model is established based on segmented Markov chains. It considers the sparseness of the distribution of the CAN network errors. Second, by learning the differences between the model estimates and the actual values from the observable node, a Bayesian network is developed for the estimation updating mechanism of the observable nodes. Then, this estimation updating mechanism is transferred to general CAN nodes with no TEC value accessibility to update the TEC estimation. Finally, a node reliability assessment method is developed to predict the time to reach bus-off state of the nodes. Case studies are carried out to demonstrate the effectiveness of the proposed methodology. Experimental results show that the estimates using the proposed model agree well with actual observations.

**Key words:** Controller area network (CAN); Transmit error counter (TEC); TEC value estimation; Bayesian network; Bus-off hitting time

http://dx.doi.org/10.1631/FITEE.1601029 **CLC number:** TP277

## 1 Introduction

Controller area network (CAN) is one of the most widely used fieldbuses in industrial automation. Developed in the 1980s, CAN networks provide good real-time performance and reliability at relatively low cost. With the growing complexity of networked automation systems and the diversity of industrial conditions, the reliability of the CAN bus network decreases gradually over time (Cauffriez *et al.*, 2003). In practice, factors such as hardware dysfunction and electromagnetic interference will result in communication delay or interruption and, more severely, may lead to the failure of the network, which will be catastrophic in actual industrial applications. To ensure production safety and product quality, higher requirements for the performance and reliability of the network should be satisfied. To improve the performance and reliability of the system, and reduce maintenance costs, it becomes increasingly important to evaluate CAN network

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (Nos. 51475422 and 51521064) and the National Basic Research Program (973) of China (No. 2013CB035405)

# A preliminary version was presented at the IEEE International Conference on Automation Science and Engineering (CASE), Aug. 24–28, 2015, Gothenburg, Sweden

ORCID: Lei-ming ZHANG, http://orcid.org/0000-0002-6002-8852

reliability online, which requires to estimate the time to reach the bus-off state of the nodes in the network.

In the literature, some studies focused on network performance and reliability. For example, Zhao and Lei (2012) and Lei *et al.* (2014) classified the error frames on the bus according to the length of the error flag, and then combined a generalized zero-inflated Poisson (GZIP) model with the ranked probability control (RPC) charts to achieve failure early warning. Navet *et al.* (2000) established a generalized Poisson process (GPP) model to describe both error frequency and error gravity, and evaluated the worst-case deadline failure in real-time applications. Wang *et al.* (2010) designed a fault-tolerant redundancy technology to enhance the reliability of the CAN bus, which integrates the advantages of the analytical and hardware redundancy fault-tolerant control concepts. Barranco *et al.* (2006; 2011) introduced an active star topology, i.e., CANcentrate. The active hub can prevent error propagation from any of its ports to the others, and has a better fault tolerance performance. Yomsi *et al.* (2012) proposed a framework based on the transaction model, and integrated offsets in the analysis of worst-case response times (WCRT) on CAN. Kumar *et al.* (2009) discussed the response time model of CAN messages, and introduced a technique to derive probabilistic parameters for messages with deterministic transmission times.

In addition, some studies focused on predicting the time of the nodes reaching the bus-off state, which also involves modeling on the variation of a node's embedded transmit error counter (TEC) value. For example, Gaujal and Navet (2005) derived a continuous time Markov chain of the CAN fault confinement mechanism, and assessed the risk of reaching the error passive or bus-off state. Lei *et al.* (2010) analyzed the changing trend of the CAN node TEC value, established a discrete time Markov chain, and developed a network/node time to failure prediction algorithm. Chen *et al.* (2006) set up a discrete time Markov model of the TEC value change, and then calculated how long a CAN node can work normally before it goes into its bus-off state. Navet and Song (2001) modeled the TEC value change by a discrete time Markov model under three cases: no message is sent, send a message successfully, and fail to send a message, and predicted the hitting time of bus-off for nodes. These studies are based on the bit error rate (BER) on the bus. However, accurate measurement of the BER is difficult, which limits the application of the above methods.

According to the literature, to analyze the time to reach the bus-off state of a node, existing methods need to measure the BER on the bus, which is an averaged measure that could not reflect the stochastic behaviors of the nodes upon each error. Moreover, existing bus-off hitting time estimation methods assume that the current internal error counters are zero due to the inaccessibility of the node internal states. Hence, the estimation results are not accurate when the errors are persistent in the network. In addition, although some intelligent nodes are commercially available to provide network error information, this information cannot be used to evaluate the reliability of the whole network system, especially when traditional nodes are also used in the network. Therefore, it is needed to develop a new node reliability assessment methodology that can predict the node bus-off hitting time without interrupting normal network communication.

The goal of this study is to develop a novel node reliability assessment methodology for a CAN network using network traffic records and information from TEC accessible nodes. The advantage of the proposed method is that the node bus-off hitting time can be predicted based on current estimates of the node TEC values for a higher reliability assessment accuracy, which is not available in existing methods. The result of this work will enable one to estimate all the nodes' statuses and time to bus-off for node reliability assessment, which will ultimately improve system maintenance practice for product quality and production safety assurance.

## 2 CAN fault confinement mechanism

The CAN fault confinement mechanism can enhance the efficiency of communication by preventing an abnormal node from disturbing the normal operation of the whole system. In the process of CAN bus communication, a node will be in one of the three states (Bosch, 1991): error active state, error passive state, and bus-off state.

The status of a CAN node is determined by its embedded TEC and receive error counter (REC). In general, a node will be in the error active state if its TEC value is less than 127. If the node TEC value is

greater than 127, the node will be in the error passive state. The node will be in the bus-off state if its TEC value is greater than 255. These counters are modified according to a set of rules (Bosch, 1991), and the simplified rules are as follows: if a node detects an error in the data received, it will send the error frames to interrupt this transmission, and the node TEC value will increase by eight, until it crosses 255. If the message is transmitted or received successfully, the node TEC or REC value will decrease by one, until it reaches zero (Farsi *et al.*, 1999; Gaujal and Navet, 2005; Chen *et al.*, 2006).

## 3 Problem statement

Given a history of network traffic, i.e., datalink and physical layers data including incomplete data frame interrupted by errors, how to accurately estimate the time to reach the bus-off state of all the nodes is a challenging task since the majority of the nodes' embedded TEC values cannot be accessed. Hence, to assess the reliabilities of all the nodes in the network, the following issues should be addressed:

1. How to establish a node TEC value estimation model according to the error log? In addition, how to deal with the case in which the data frames interrupted by the errors are unrecognizable?

2. How to establish a knowledge-based model to describe the differences between the actual TEC value and the estimated TEC value of an observable node, so that the model can be transferrable to general nodes without TEC accessibility to update the predicted TEC values based on network error information?

3. How to develop a model to predict the time to reach the bus-off state of a node with the knowledge of the node updated TEC value considering the sparse distribution of the errors in time?

The assumptions in this study are as follows: (1) only one master device (PLC) is in the network, and the communication sets as polling mode; (2) the TEC values of the observable nodes can be accessed without significant delay; (3) the identifier (ID) of the node can be identified from the frame ID.

## 4 Proposed methodology

In this work, a novel node reliability assessment methodology is proposed for CAN network using error records and information from TEC accessible nodes. The principal idea of the proposed method is that by establishing a data-driven model to describe the differences between the observed and estimated TEC values of an observable node from network traffic data, the TEC values of general nodes without TEC accessibility can be better estimated using a TEC estimation updating mechanism by transferring the knowledge from the model on the observable nodes.

The overall framework of the proposed method is shown in Fig. 1. First, error identification is conducted to determine the ID of the transmitting node upon each error. Second, considering the sparseness of the errors, the node TEC value can be estimated by segmented discrete time Markov chains (Zhang *et al.*, 2015). Third, a Bayesian network is built to learn the difference between the estimated and actual TEC values of the observable node. Finally, based on the updated TEC value using the Bayesian network transferred to general nodes, the random walk model is developed to predict the node bus-off hitting time. Details of the method are introduced as follows.

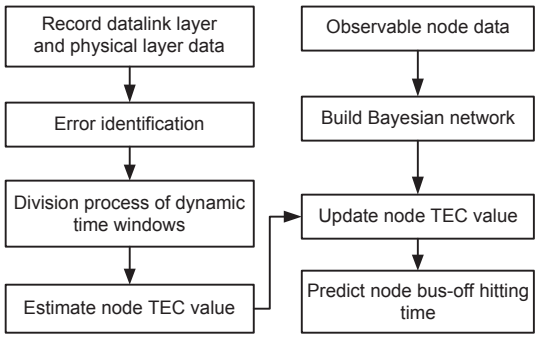


**Fig. 1 Overall framework of the proposed method (TEC: transmit error counter)**

### 4.1 Node TEC value estimation

#### 4.1.1 Error identification

Error identification is a procedure to identify which node is transmitting when a data frame is interrupted by an error, so that the corresponding node TEC value can be calculated. Fig. 2 illustrates the scenario where a data frame is interrupted by an error. According to the CAN protocol, the occurrence of an error will cause an error trigger, and then the transmission process will be interrupted by error frames.

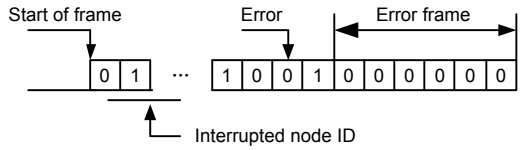


**Fig. 2 Data frame interrupted by an error**

Each CAN data frame has an arbitration field which contains the information of the node that is transmitting this message. If the ID information of the interrupted data frame is complete, it can be directly extracted from the interrupted data frame, and then the error frame can be categorized to the transmitting node. However, due to the implementation of the CAN protocol chips, an incomplete CAN data frame will be discarded. Specific FPGA based hardware and software systems are developed so that all the incomplete data frames upon each error can be recorded using the error triggering and recording mechanism developed in Lei and Djurdjanovic (2010). The procedure of error identification is shown in Fig. 3.

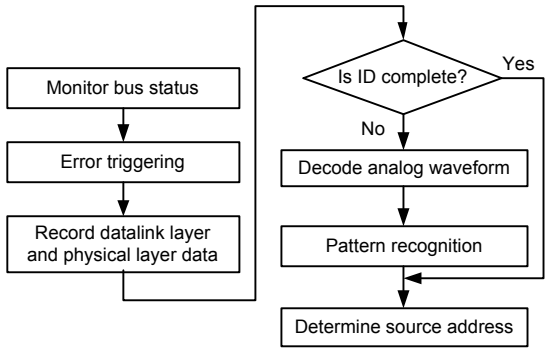


**Fig. 3 Procedure of error identification**

If the ID field of the interrupted frame is partially damaged, analog waveforms will be used to extract the source address of the interrupted packet. First, waveforms are recorded by data acquisition equipment and decoded to obtain the address segment. Second, the integrity of the address segment should be analyzed. If the address segment is partially available, the source address of the interrupted packet can be determined by conducting pattern recognition analysis. Details of the analysis procedure are given in Lei *et al.* (2010).

#### 4.1.2 Segmented discrete time Markov chains

From the recorded information, the distribution of the error events can be analyzed in the time based on their time stamps. The distribution of error frames over time is illustrated in Fig. 4, in which the arrows represent the occurrence of the errors. In general, the number of error frames is low over the time period; i.e., the network errors occur only sporadically. Also, the inter-arrival intervals of the errors may not follow the same distribution at different time periods.

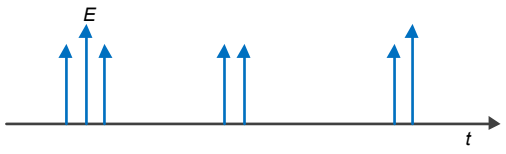


**Fig. 4 Distribution of error frames in the time axis**

Considering the error distribution fluctuation in TEC value estimation, dynamic time windows are used to segment the time axis, making the distribution of error frames in each time window more coherent. The length of time windows is unequal, and changes dynamically according to the distribution of the error events along the time axis (Fig. 5).

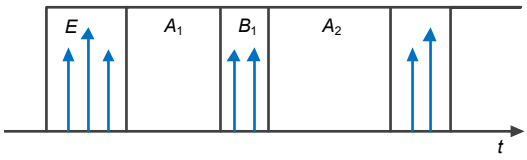


**Fig. 5 Division of dynamic time windows**

The division process of the dynamic time windows is shown in Fig. 6. Assume that the average number of successful transmissions between two error frames is $t_{\mathrm{trsb}}$, and that the standard deviation for the number of successful transmissions is $s_{\mathrm{d}}$. If the number of successful transmissions $n$ between two error frames is greater than $t_{\mathrm{trsb}} + s_{\mathrm{d}}$, a time window should be put between the two error frames (e.g., the blank windows $A_1$ and $A_2$ in Fig. 5). Then the part between the two blank time windows is the division for another time window, which contains error frames, such as $B_1$ in Fig. 5.

The successful transmission probability within each time window can be obtained by

$$p_i = \mathrm{St}_i/(\mathrm{Err}_i + \mathrm{St}_i), \tag{1}$$

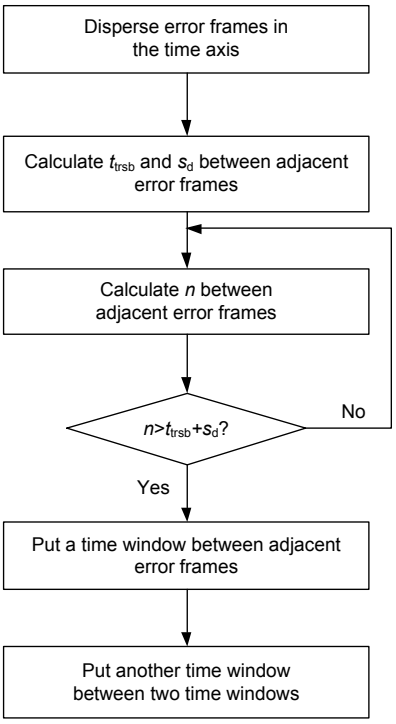


**Fig. 6 Division process of dynamic time windows**

where $p_i$ is the successful transmission probability in the $i$th time window, $\mathrm{St}_i$ denotes the number of successful transmissions, and $\mathrm{Err}_i$ the number of error frames in the $i$th time window. The node TEC value one-step transition probability matrix $\boldsymbol{P}_i$ of the Markov chain in the $i$th time window is shown in Eq. (2) (at the bottom of this page). In this matrix, the first column denotes the current TEC value, and the first row denotes the next value.

Assuming the initial distribution of the TEC value is

$$\boldsymbol{\alpha} = [1, 0, \ldots, 0]^{\mathrm{T}} \in \mathbb{R}^{257}, \tag{3}$$

each possible value for TEC is

$$\boldsymbol{Z} = [0, 1, \ldots, 256]^{\mathrm{T}} \in \mathbb{R}^{257}, \tag{4}$$

and the expectation of the estimated TEC value can be obtained by

$$E[\mathrm{TEC}] = \boldsymbol{\alpha}^{\mathrm{T}} \cdot \boldsymbol{P}_1^{N_1} \cdot \boldsymbol{P}_2^{N_2} \cdots \boldsymbol{P}_n^{N_n} \cdot \boldsymbol{Z}, \tag{5}$$

where $n$ is the number of time windows, and $N_i$ the number of node TEC value changes in the $i$th time window, $N_i = \mathrm{St}_i + \mathrm{Err}_i$.

The variance of the estimated TEC value can be obtained by

$$D[\mathrm{TEC}] = \boldsymbol{\alpha}^{\mathrm{T}} \cdot \boldsymbol{P}_1^{N_1} \cdot \boldsymbol{P}_2^{N_2} \cdots \boldsymbol{P}_n^{N_n} \cdot \boldsymbol{Z}^2 - (E[\mathrm{TEC}])^2, \tag{6}$$

where $\boldsymbol{Z}^2$ is making a square operation for each element of $\boldsymbol{Z}$.

4.1.3 TEC value estimation considering unidentified error information

If an error occurs in the arbitration area of a data frame, the node ID of the corrupted data frame cannot be identified by using only datalink layer information. In this case, node ID can be identified directly using pattern recognition analysis from the physical layer signals (Lei and Djurdjanovic, 2010). However, a special scenario may occur when no useful data is recorded. Hence, two extreme approaches are proposed to deal with this situation:

**Approach 1** None of the unrecognizable data frames is related to the estimated node.

**Approach 2** All of unrecognizable data frames are classified to the estimated node.

The comparison of the effects of these two approaches on the TEC value estimation will be discussed in the experimental study section.

$$\boldsymbol{P}_i = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \cdot \\ \cdot \\ \cdot \\ 254 \\ 255 \\ 256 \end{array} \begin{array}{c} \begin{array}{ccccccccccc} 0 & 1 & 2 & \cdots & 8 & 9 & \cdots & 253 & 254 & 255 & 256 \end{array} \\ \begin{bmatrix} p_i & 0 & 0 & \cdots & 1-p_i & 0 & \cdots & 0 & 0 & 0 & 0 \\ p_i & 0 & 0 & \cdots & 0 & 1-p_i & \cdots & 0 & 0 & 0 & 0 \\ 0 & p_i & 0 & \cdots & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & p_i & 0 & 0 & 1-p_i \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & p_i & 0 & 1-p_i \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}. \tag{2}$$

### 4.2 Updating on the node TEC value

The assumption of the node initial TEC value distribution in Eq. (2) is that the node TEC value is zero when we start estimation. However, in practice, the estimation does not always begin from the boot time. In addition, the initial distribution of the TEC value is not precisely known. Hence, there exists deviations between the estimated TEC value and the actual value. Therefore, it is necessary to update the estimated TEC value based on the information learned from the observable nodes.

#### 4.2.1 TEC value updating procedure

The proposed updating procedure of the node TEC value is shown in Fig. 7.

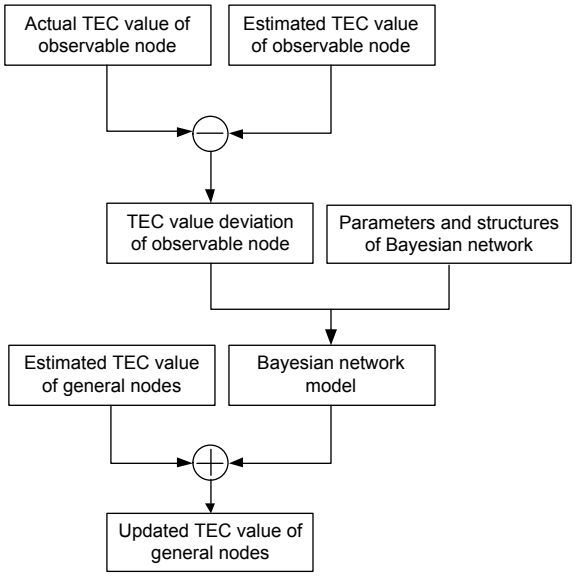


**Fig. 7 Procedure of updating the node TEC value**

Let $\xi$ be the deviation of a node TEC value between the actual value $Z_{\mathrm{a}}$ and the estimated value $Z_{\mathrm{e}}$ described in Eq. (7):

$$Z_{\mathrm{a}} = Z_{\mathrm{e}} + \xi. \tag{7}$$

Since the deviations of all nodes' estimated TEC values are the results of the same set of variables, they can be modeled by the same Bayesian network.

On one hand, for an observable node whose actual TEC value $Z_{\mathrm{a}}$ is accessible, the estimated TEC value $Z_{\mathrm{e}}$ using network traffic data can be obtained by Eq. (5). Then, the deviation $\xi$ can be calculated by Eq. (7), which can be used to conduct the parameter study of the Bayesian network. On the other hand, for general nodes, the estimated TEC value $Z_{\mathrm{e}}$ can be calculated by Eq. (5), and then updated after obtaining the correcting deviation $\xi$ from the Bayesian network for this node.

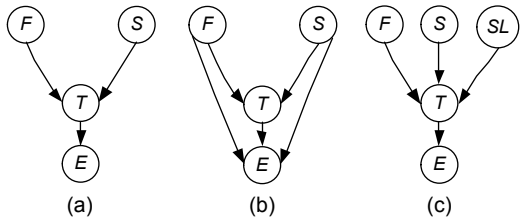


**Fig. 8 Bayesian network structures in this study: (a) structure 1; (b) structure 2; (c) structure 3**

#### 4.2.2 Bayesian network construction

Three Bayesian network structures are constructed in this section. In structure 1 (Fig. 8a), the number of errors ($F$) and the number of successful transmissions ($S$) in the last time window are the parent nodes of the estimated TEC value ($T$). The deviation ($E$) is the child node of $T$. Compared with structure 1, structure 2 (Fig. 8b) establishes two direct links, i.e., $F$ to $E$ and $S$ to $E$. Compared with structure 1, a parent node is added in structure 3 (Fig. 8c), which is the length of the second to last time window ($SL$).

Under the condition of different variables, the conditional expectation of deviation $E[D|X]$ can be obtained by the Bayesian network:

$$E[D|X] = \sum_{i=1}^{n} d_i P(D = d_i|X), \tag{8}$$

where $d_i$ is the $i$th value of the deviation, $X \in \{F, S, T\}$ within structures 1 and 2, and $X \in \{F, S, SL, T\}$ within structure 3.

Hence, the estimated TEC value can be updated with the conditional expectation of deviation:

$$Z_{\mathrm{u}} = Z_{\mathrm{e}} + E[D|X], \tag{9}$$

where $Z_{\mathrm{u}}$ denotes the updated TEC value.

### 4.3 Node bus-off hitting time prediction

In this section, the random walk model is developed to describe the dynamics of the node TEC value and estimate the number of steps to be taken until a node goes into its bus-off state. The node bus-off hitting time is calculated with the knowledge of the random walk model and the updated TEC value.

$$N_0 = \begin{cases} 1 + N_0, & \text{with probability } p, \\ 1 + N_8, & \text{with probability } 1 - p, \end{cases}$$
$$N_{k,\ 1\le k\le 247} = \begin{cases} 1 + N_{k-1}, & \text{with probability } p, \\ 1 + N_{k+8}, & \text{with probability } 1 - p, \end{cases} \tag{10}$$
$$N_{k,\ 248\le k\le 256} = \begin{cases} 1 + N_{k-1}, & \text{with probability } p, \\ 1, & \text{with probability } 1 - p. \end{cases}$$

$$E[N_k] = \begin{cases} p \cdot E[1 + N_0] + (1 - p) \cdot E[1 + N_8], & k = 0, \\ p \cdot E[1 + N_{k-1}] + (1 - p) \cdot E[1 + N_{k+8}], & 1 \le k \le 247, \\ p \cdot E[1 + N_{k-1}] + (1 - p) \cdot E[1 + N_{256}], & 248 \le k < 255, \\ 0, & k = 256. \end{cases} \tag{11}$$

4.3.1 Random walk model

The change of the node TEC value in Eq. (2) is the random walk process (Janssen, 1981). Assume that a node's current TEC value is $k$, and $N_k$ denotes the number of steps taken such that a node TEC value changes from $k$ to 256. Using 'one-step' analysis (Gaujal and Navet, 2005), $N_k$ can be obtained by Eq. (10). Further, the expectation of $N_k$ can be obtained by Eq. (11).

In our work, the communication mode acts as a polling mode. Hence, the expectation of node bus-off hitting time $T_{\text{hit}}$ is proportional to the number of successful transmissions during the process:

$$T_{\text{hit}} = p \cdot E[N_k] \cdot T_{\text{pol}}, \tag{12}$$

where $T_{\text{pol}}$ denotes the polling cycle, and $p$ the probability of successful transmission. The standard deviation of node bus-off hitting time $T_\sigma$ is obtained as

$$T_\sigma = p \cdot \sqrt{D[N_k]} \cdot T_{\text{pol}}, \tag{13}$$

where $D[N_k]$ is the variance of $N_k$, which can be obtained as

$$D[N_k] = E[N_k^2] - (E[N_k])^2. \tag{14}$$

4.3.2 Successful transmission probability calculation

We assume that the number of TEC value changes within a polling cycle $X$ follows a geometric model:

$$P(X = x) = (1 - p)^{x-1} \cdot p, \tag{15}$$

where $p$ can be estimated by the maximum likelihood estimation (MLE) algorithm using the following equation:

$$\hat{p} = \frac{n}{\sum_{i=1}^{n} X_i}, \tag{16}$$

where $n$ is the number of polling cycles.

## 5 Testbed setup and case studies

In this section, we first introduce the details of the testbed. Second, case studies are carried out to verify the effectiveness of the node TEC value estimation model. Third, the update of the node estimated TEC value is realized by the Bayesian network. In addition, the optimal Bayesian network structure is selected based on the results of the case studies. Finally, the node bus-off hitting time is calculated for different cases, and the effectiveness of prediction is verified.

### 5.1 Testbed setup

The schematic layout of the experiment testbed is shown in Fig. 9, and the constructed testbed is shown in Fig. 10. The testbed contains three parts: communication platform, error injection system, and data acquisition system. The communication platform is constructed by DeviceNet protocol hardware, which is based on the CAN protocol. In the following case studies, the bus communication speed is set at 500 kb/s, and there are five nodes on the bus, i.e., nodes 8, 9, 12, 15, and the PLC. Nodes 12 and 15 are in-house developed observable nodes whose TEC value and time stamp of the TEC increments can be accessed and transmitted.

The error injection system injects network errors using a high-speed switch to emulate the intermittent

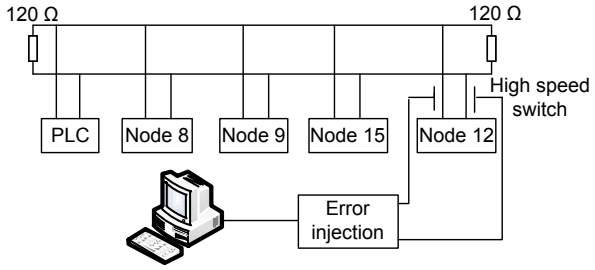


**Fig. 9 Schematic layout of the experimental network**

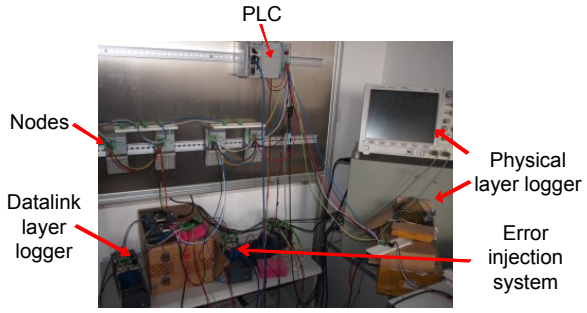


**Fig. 10 Experimental testbed for case studies (References to color refer to the online version of this figure)**

connection problems on the network cables. The open circuit of the high-speed switch is controlled by the LabVIEW programs running on the FPGA, which creates the transient disconnection events of CAN cables, and results in the network errors. The disconnection events' arrivals follow a Poisson distribution. In this study, the default average fault injection interval is 800 μs on the drop cable of node 12, unless otherwise specified. Note that although the disconnection event is Poisson, the resultant occurrence of the errors is not Poisson (Lei *et al.*, 2014).

The data acquisition system includes the datalink layer logger and the physical layer logger, which collect the datalink layer information and physical layer information upon errors to predict the node bus-off hitting time.

## 5.2 TEC value estimation for observable node

Two case studies are carried out on the observable node 15. The first one verifies the effectiveness of segmented discrete time Markov chains. The second one analyzes the two approaches for dealing with the unrecognizable data frames.

### 5.2.1 TEC value estimation using segmented discrete time Markov chains

Segmented discrete time Markov chains are established to estimate the node TEC value. The comparison of the estimated TEC value with the actual observations under 800 μs error injection intervals is shown in Fig. 11.

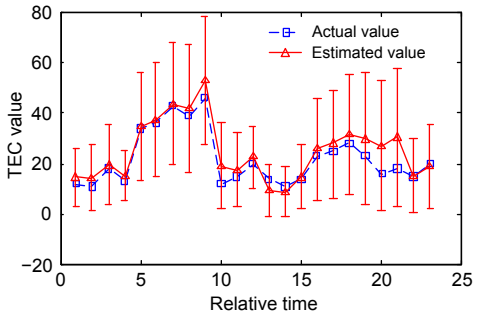


**Fig. 11 Comparison of the estimated TEC value with the actual value (800 μs)**

The upper and lower bounds of error bars in Fig. 11 are the positive and negative standard deviations of the estimated TEC value. Fig. 11 shows that the estimated TEC value agrees well with the actual value. A closer look at the figure shows that the estimated TEC value has a certain volatility, which is consistent with the actual situation.

### 5.2.2 TEC estimation with unrecognizable data frames

In this case study, unrecognizable data frames are randomly marked so that results can be compared. First, 3% of recorded interrupted data frames are randomly marked as unrecognizable. Then the segmented discrete time Markov chains are established to estimate the node TEC value using two unrecognizable data approaches. The comparisons of the estimated TEC value using data treatment approaches 1 and 2 with the actual TEC value are shown in Fig. 12.

Fig. 12 shows that the differences between estimated and actual TEC values using data treatment approach 1 are much smaller than those using approach 2. Moreover, as shown in Table 1, the estimation error norms using approach 1 are significantly smaller than those using approach 2. Therefore, approach 1 with none of the unrecognizable data frames classified to the estimated node is better.

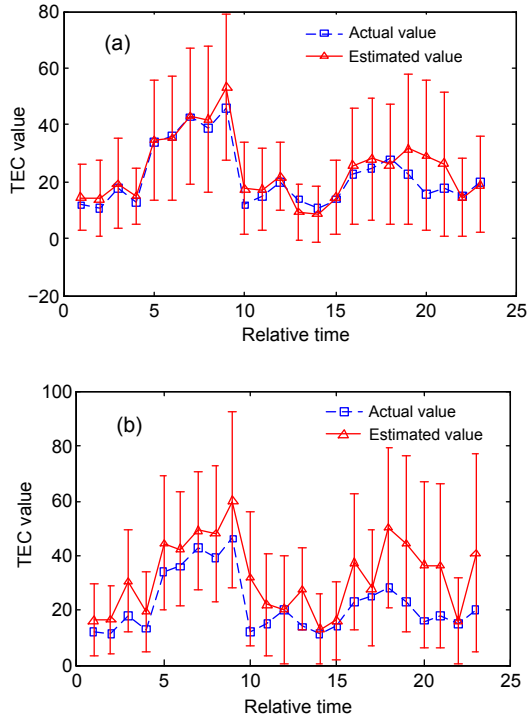


**Fig. 12 Comparison of the estimated TEC value using data treatment approach 1 (a) and approach 2 (b) with the actual TEC value**

**Table 1 Error norms of two approaches**

| Approach | 1-norm | 2-norm | $\infty$-norm |
|---|---|---|---|
| 1 | 77.87 | 22.23 | 13.15 |
| 2 | 239.84 | 60.47 | 21.98 |

Further analysis shows that the unrecognizable data frames are related not only to the estimated node, but also to the other nodes interrupted by the errors. In our testbed, normally eight data frames are transmitted in each polling cycle, and hence if the probability of the data frame being interrupted in the arbitration area is equal among different nodes, the unrecognizable data frames related to the estimated node account only for 1/8 of all the unrecognizable data frames. The larger the number of nodes, the lower the probability that the unrecognizable data frames belong to the estimated node. Hence, in the case where the interrupted data frames cannot be identified only by their ID information, using approach 1 is suitable for node TEC estimation.

### 5.3 Bayesian network and optimal structure selection

First, the data from observable node 15 are used to train the Bayesian network structures. Then, these Bayesian networks are transferred to observable node 12 to demonstrate the effectiveness of the TEC value updating procedure by comparing the updated TEC values and the actual values. The settings of various conditions are shown in Table 2. The conditions in each case study denote which variables are used in the conditional expectation; e.g., $T$ means only $T$ is used, and $TF$ means both $T$ and $F$ variables are used.

**Table 2 Case studies settings for Bayesian networks based updating**

| Case | Bayesian structure | Conditions |
|---|---|---|
| 1 | Fig. 8a | $T$, $F$, $S$, $TF$, $TS$, $FS$, $TFS$ |
| 2 | Fig. 8b | $T$, $F$, $S$, $TF$, $TS$, $FS$, $TFS$ |
| 3 | Fig. 8c | $T$, $F$, $S$, $SL$, $TF$, $TS$, $FS$, $TFS$ |

#### 5.3.1 Effectiveness of Bayesian network structure 1

The update results obtained with conditions $T$, $F$, and $S$ when applying the Bayesian network shown in Fig. 8a, are shown in Fig. 13. The norms of the updated TEC value error with the three different conditions are shown in Table 3.

Table 3 shows that the 1- and 2-norm of the updated TEC value error with conditions $T$, $F$, and $S$ decrease significantly when compared to the corresponding norms of the estimated TEC value error. The $\infty$-norm of the updated TEC value error with condition $S$ also decreases. Hence, the updating effect with condition $T$ is more significant than that with $F$ or $S$.

The updating results with conditions $TF$, $TS$, $FS$, and $TFS$ are shown in Fig. 14. The norms of the updated TEC value error with the four different conditions are also shown in Table 3.

Table 3 shows that the 1- and 2-norm of the updated TEC value error with conditions $TF$, $TS$, $FS$, and $TFS$ also decrease when compared to the norms of the estimated TEC value error, but the decreasing tendency is not significant. In addition, the $\infty$-norm of the updated TEC value error remains at the same level. However, from Table 3 , the update effects with conditions $T$, $F$, and $S$ are more significant than those with conditions $TF$, $TS$, $FS$, and $TFS$. Therefore, only conditions $T$, $F$, and $S$ need to be considered when applying Bayesian network structure 1.

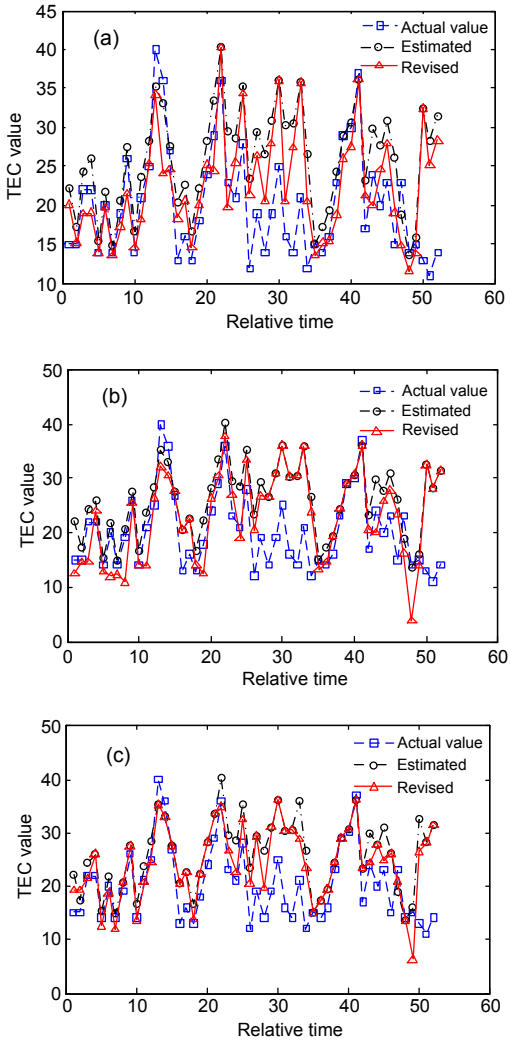


**Fig. 13 Updating results with conditions *T* (a), *F* (b), and *S* (c) in structure 1**

**Table 3 Norms of the updated TEC value error with different conditions in structure 1**

| Condition | 1-norm | 2-norm | ∞-norm |
|---|---|---|---|
| Estimated TEC value error | 322.50 | 58.38 | 19.49 |
| *T* | 260.75 | 48.11 | 19.49 |
| *F* | 309.19 | 56.92 | 19.49 |
| *S* | 263.39 | 50.58 | 17.39 |
| *TF* | 302.17 | 57.08 | 19.49 |
| *TS* | 317.08 | 58.28 | 19.49 |
| *FS* | 303.48 | 57.17 | 19.49 |
| *TFS* | 317.08 | 58.28 | 19.49 |

### 5.3.2 Effectiveness of Bayesian network structure 2

The updating results with different conditions listed in Table 2 can be obtained when applying the Bayesian network in Fig. 8b. The norms of the updated TEC value error with different conditions are shown in Table 4.

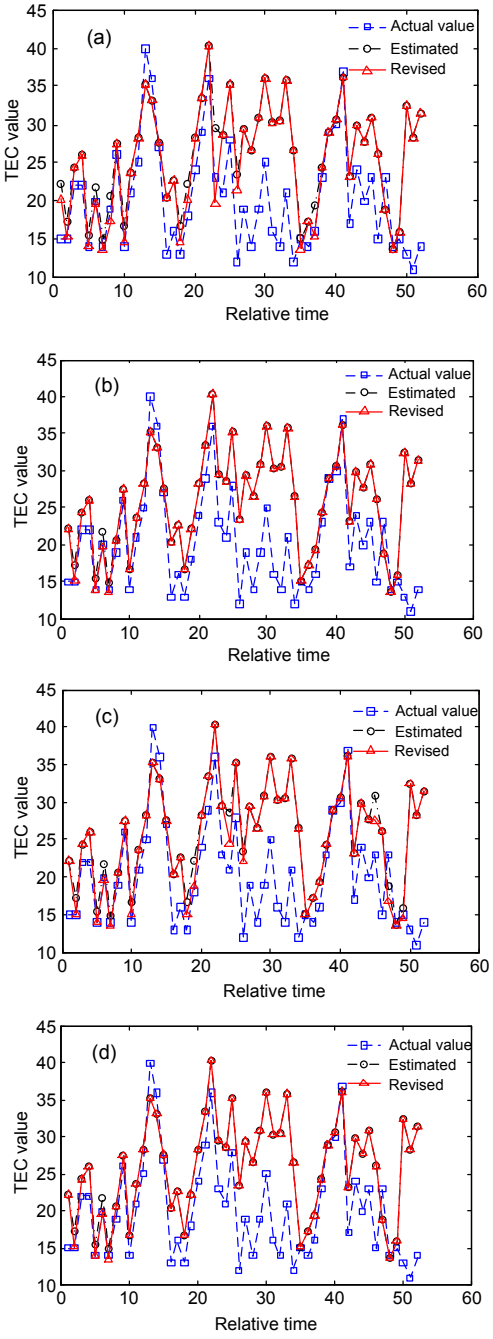


**Fig. 14 Updating results with conditions *TF* (a), *TS* (b), *FS* (c), and *TFS* (d) in structure 1**

Table 4 shows that the 1- and 2-norm of the updated TEC value error with conditions *T*, *TF*, *TS*, *FS*, and *TFS* decrease when compared to the norms of the estimated TEC value error, but the effect is not obvious, and the ∞-norm of the updated TEC value error has not changed. Moreover, with the conditions

**Table 4 Norms of the updated TEC value error with different conditions in structure 2**

| Condition | 1-norm | 2-norm | ∞-norm |
|---|---|---|---|
| Estimated TEC value error | 322.50 | 58.38 | 19.49 |
| *T* | 261.67 | 48.75 | 19.49 |
| *F* | 322.50 | 58.38 | 19.49 |
| *S* | 322.50 | 58.38 | 19.49 |
| *TF* | 306.80 | 57.39 | 19.49 |
| *TS* | 317.08 | 58.29 | 19.49 |
| *FS* | 307.75 | 57.65 | 19.49 |
| *TFS* | 317.08 | 58.29 | 19.49 |

*F* and *S*, none of the error norms decreases. Hence, these two conditions have no effect.

Compared with the updating effect of Bayesian network structure 1, although structure 2 has a certain effect, structure 1 is more effective. A similar conclusion can be drawn from comparing Bayesian network structures 1 and 3, whose results are omitted due to page limits.

#### 5.3.3 Optimal structure selection

The above analysis shows that Bayesian network structure 1 shows the best updating effects. In addition, structures 2 and 3 are much more complex, which lead to greater computation and lower efficiency. Therefore, Bayesian network structure 1 should be used to update the node estimated TEC value, and only conditions *T*, *F*, and *S* need to be considered.

### 5.4 Bus-off hitting time prediction for general nodes

In this subsection, the average fault injection interval is set as 800, 810, and 820 μs, and the fault is injected on the drop cable on node 9. Assume $t_p$ is half of the node actual bus-off hitting time, and the data recorded before $t_p$ are used to predict the node bus-off hitting time. First, the estimated TEC value is obtained by segmented discrete time Markov chains using network traffic data. Then, the updating of the node estimated TEC is realized by Bayesian network structure 1 (Fig. 8a), which is trained using data from observable node 12. Finally, the node bus-off hitting time is calculated by the random walk model and compared with actual observation.

Two sets of case studies are carried out. The first one predicts the node bus-off hitting time under different error log deletion ratios, used to emulate a random un-identifiable data frame scenario caused by errors. The second study compares the proposed method with the traditional TEC value using Markov chains developed in Lei *et al.* (2010).

#### 5.4.1 Bus-off time prediction under different error log deletion ratios

Tables 5 gives the corresponding prediction of node bus-off hitting times for the error log deletion ratios equal to 25%, 50%, and 75%. The predicted node bus-off hitting time agrees well with the actual value under different error log deletion ratios.

#### 5.4.2 Hitting time prediction using Markov chain

As mentioned in Section 4.3.1, let $N_k$ denote the number of steps that a node TEC value changes from $k$ to 256. Here, a Markov chain model, developed by Lei *et al.* (2010), can be used to calculate the expectation and variance of $N_k$. Then, the expectation and the standard deviation of node bus-off hitting time can be obtained by Eqs. (12) and (13), respectively. The prediction results are shown in Table 6. The deviations between the predicted value and actual value of node bus-off hitting time using the Markov chain model are significantly larger than that using the method proposed in this study. Hence, the proposed method shows better prediction results.

**Table 5 Comparison of predicted bus-off time with actual time for different error log deletion ratios**

| Average fault injection interval (μs) | Actual hitting time (s) | | | Expectation of predicted hitting time (s) | | | Standard deviation of predicted hitting time (s) | | |
|---|---|---|---|---|---|---|---|---|---|
| | $R_E$=25% | $R_E$=50% | $R_E$=75% | $R_E$=25% | $R_E$=50% | $R_E$=75% | $R_E$=25% | $R_E$=50% | $R_E$=75% |
| 800 | 16.79 | 16.79 | 16.79 | 15.25 | 14.95 | 10.60 | 6.90 | 6.70 | 4.12 |
| 810 | 67.27 | 67.27 | 67.27 | 80.41 | 68.73 | 54.55 | 64.87 | 53.56 | 40.05 |
| 820 | 95.89 | 95.89 | 95.89 | 106.53 | 79.65 | 64.88 | 90.49 | 64.12 | 48.87 |

$R_E$: error log deletion ratio

**Table 6 Comparison of predicted bus-off time and actual time (Markov chain model)**

| Average fault injection interval (μs) | Actual hitting time (s) | Expectation of predicted hitting time (s) | Standard deviation of predicted hitting time (s) |
|---|---|---|---|
| 800 | 16.79 | 22.27 | 11.83 |
| 810 | 67.27 | 79.15 | 63.64 |
| 820 | 95.89 | 117.16 | 100.99 |

## 6 Conclusions

In this paper, a CAN node reliability assessment methodology has been proposed using the information from observable nodes. First, the node TEC value is estimated by a segment discrete time Markov chain using the data recorded on the bus. In addition, two approaches are proposed and compared to handle the unrecognizable data frames caused by a damaged ID segment. Then, the optimal Bayesian network is trained using the TEC values from the observable nodes, and a TEC value updating procedure for general nodes is proposed. Finally, the random walk model is established to predict the node bus-off hitting time. Case studies are carried out to demonstrate the effectiveness of the proposed methodology. Experimental results show that the estimated TEC values agree well with the actual observations, as well as the predicted bus-off hitting time of the nodes. Future work includes optimizing the method of time window division, improving TEC value updating procedures, and dealing with the multi-master mode.

## References

Barranco, M., Proenza, J., Rodríguez-Navas, G., *et al.*, 2006. An active star topology for improving fault confinement in CAN networks. *IEEE Trans. Ind. Inform.*, **2**(2):78-85. http://dx.doi.org/10.1109/TII.2006.875505

Barranco, M., Proenza, J., Almeida, L., 2011. Quantitative comparison of the error-containment capabilities of a bus and a star topology in CAN networks. *IEEE Trans. Ind. Electron.*, **58**(3):802-813. http://dx.doi.org/10.1109/TIE.2009.2036642

Bosch, 1991. CAN Specification Version 2.0. Robert Bosch GmbH, Postfach, Germany.

Cauffriez, L., Conrard, B., Thiriet, J., *et al.*, 2003. Fieldbuses and their influence on dependability. Proc. 20th IEEE Instrumentation and Measurement Technology Conf., p.1005-1008. http://dx.doi.org/10.1109/IMTC.2003.1208126

Chen, J.X., Luo, F., Sun, Z.C., 2006. Reliability analysis of CAN nodes under electromagnetic interference. IEEE Int. Conf. on Vehicular Electronics and Safety, p.367-371. http://dx.doi.org/10.1109/ICVES.2006.371617

Farsi, M., Ratcliff, K., Barbosa, M., 1999. An overview of controller area network. *Comput. Contr. Eng. J.*, **10**(3):113-120. http://dx.doi.org/10.1049/cce:19990304

Gaujal, B., Navet, N., 2005. Fault confinement mechanisms on CAN: analysis and improvements. *IEEE Trans. Veh. Technol.*, **54**(3):1103-1113. http://dx.doi.org/10.1109/TVT.2005.844652

Janssen, H.K., 1981. On the nonequilibrium phase transition in reaction-diffusion systems with an absorbing stationary state. *Zeitschr. Phys. B*, **42**(2):151-154. http://dx.doi.org/10.1007/BF01319549

Kumar, M., Verma, A.K., Srividya, A., 2009. Response-time modeling of controller area network (CAN). Int. Conf. on Distributed Computing and Networking, p.163-174. http://dx.doi.org/10.1007/978-3-540-92295-7_20

Lei, Y., Djurdjanovic, D., 2010. Diagnosis of intermittent connections for DeviceNet. *Chin. J. Mech. Eng.*, **23**(5):606-612. http://dx.doi.org/10.3901/CJME.2010.05.606

Lei, Y., Djurdjanovic, D., Ni, J., 2010. DeviceNet reliability assessment using physical and data link layer parameters. *Qual. Reliab. Eng. Int.*, **26**(7):703-715. http://dx.doi.org/10.1002/qre.1131

Lei, Y., Yuan, Y., Zhao, J.Z., 2014. Model-based detection and monitoring of the intermittent connections for CAN networks. *IEEE Trans. Ind. Electron.*, **61**(6):2912-2921. http://dx.doi.org/10.1109/TIE.2013.2272277

Navet, N., Song, Y.Q., 2001. Validation of in-vehicle real-time applications. *Comput. Ind.*, **46**(2):107-122. http://dx.doi.org/10.1016/S0166-3615(01)00123-3

Navet, N., Song, Y.Q., Simonot, F., 2000. Worst-case deadline failure probability in real-time applications distributed over controller area network. *J. Syst. Arch.*, **46**(7):607-617. http://dx.doi.org/10.1016/S1383-7621(99)00016-8

Wang, Z.Y., Guo, X.S., Yu, C.Q., 2010. Research of fault-tolerant redundancy and fault diagnosis technology based on CAN. 2nd Int. Conf. on Advanced Computer Control, p.287-291. http://dx.doi.org/10.1109/ICACC.2010.5487002

Yomsi, P.M., Bertrand, D., Navet, N., *et al.*, 2012. Controller area network (CAN): response time analysis with offsets. 9th IEEE Int. Workshop on Factory Communication Systems, p.43-52. http://dx.doi.org/10.1109/WFCS.2012.6242539

Zhang, L.M., Tang, L.H., Yang, F., *et al.* 2015. CAN node reliability assessment using segmented discrete time Markov chains. IEEE Int. Conf. on Automation Science and Engineering, p.231-236. http://dx.doi.org/10.1109/CoASE.2015.7294067

Zhao, J.Z., Lei, Y., 2012. Modeling for early fault detection of intermittent connections on controller area networks. IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics, p.1135-1140. http://dx.doi.org/10.1109/AIM.2012.6265905