DOI: 10.3901/CJME.2014.0527.301, available online at www.springerlink.com; www.cjmenet.com; www.cjmenet.com.cn

# Fault Location Identification for Localized Intermittent Connection Problems on CAN Networks

LEI Yong[1, 2, *], YUAN Yong[1], and SUN Yichao[1]

1 *State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Hangzhou 310027, China*
2 *State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing 100084, China*

Received September 26, 2013; revised May 23, 2014; accepted May 27, 2014

**Abstract:** The intermittent connection(IC) of the field-bus in networked manufacturing systems is a common but hard troubleshooting network problem, which may result in system level failures or safety issues. However, there is no online IC location identification method available to detect and locate the position of the problem. To tackle this problem, a novel model based online fault location identification method for localized IC problem is proposed. First, the error event patterns are identified and classified according to different node sources in each error frame. Then generalized zero inflated Poisson process(GZIP) model for each node is established by using time stamped error event sequence. Finally, the location of the IC fault is determined by testing whether the parameters of the fitted stochastic model is statistically significant or not using the confident intervals of the estimated parameters. To illustrate the proposed method, case studies are conducted on a 3-node controller area network(CAN) test-bed, in which IC induced faults are imposed on a network drop cable using computer controlled on-off switches. The experimental results show the parameters of the GZIP model for the problematic node are statistically significant(larger than 0), and the patterns of the confident intervals of the estimated parameters are directly linked to the problematic node, which agrees with the experimental setup. The proposed online IC location identification method can successfully identify the location of the drop cable on which IC faults occurs on the CAN network.

**Keywords:** CAN network, fault location identification, GZIP model, intermittent connection

## 1 Introduction

Due to good real-time performance with low implementation and maintenance costs, CAN network has gained wide acceptance in various applications, such as distributed automation systems for automobile manufacturing systems, vehicle powertrain and safety control systems, and recently airplane sensor-actuator systems[1]. However, growing complexity of the applications in harsh environments also introduces different forms of interferences, such as electromagnetic interferences, grounding problems, bandwidth allocation problems, etc, which will lead to degraded network performance or even system level failures. Among these factors, the intermittent connection(IC) is one of common but difficult problems in practice, in the presence of transient disconnections between the field devices and the backbone network. Minor IC problem may lower the quality of service of network[2]. However severe IC problem may force nodes(stations), sometimes ones without IC problem, switch from active communication to bus-off state, which results in excessive downtimes and unnecessary maintenance since bus-off nodes will be replaced directly during maintenance. In safety critical applications, the loss of a system node will lead to system-wide halt, and hence result in unexpected downtimes. Therefore, identify the location of the IC problem is important in system maintenance. However, in traditional offline approach, locating IC problems after the system shutdown is a difficult and time consuming task since the IC problems usually cannot be repeated while the system is idle. Hence, online IC fault location identification method is highly demanded in practice.

In literature, a large number of studies were focused on network dependability. For instance, LIAN, et al[3], analyzed the performance of various networks under different structures to obtain guidance for robust controller design. TRAN, et al[4] studied the network vulnerabilities with multiple-bit error under different bit error rate conditions. ZHAO[5] developed a fault model to determine which communication link is sensitive to the interference imposed on the CAN network. RUFINO et al[6], deployed an additional CAN layer to enhance fault confinement mechanisms. From networked control perspective, HANSSON, et al[7], used the probability of missing the

* Corresponding author. E-mail: ylei@zju.edu.cn
Supported by National Natural Science Foundation of China(Grant No. 51005205), Science Fund for Creative Research Groups of National Natural Science Foundation of China(Grant No. 51221004), National Basic Research Program of China(973 Program, Grant No. 2013CB-035405), and Open Foundation of State Key Laboratory of Automotive Safety and Energy, Tsinghua University, China(Grant No. KF13011)

packet transmission deadline as a measure to indicate whether the network communication was still dependable in a noisy environment. Similar concept can be seen in Ref. [8]. Potential solution has been demonstrated by scheduling a set of messages with mixed criticality in an efficient fault-tolerant manner to ensure user specified dependability[9–10]. Moreover, the logic behaviors of networked system were studied from discrete event system perspective for fault management[11]. JIANG, et al[12], proposed an event-based method for quantitative description of repeated failures in certain time intervals. ANGSKUN, et al[13], implemented an event-based self healing mechanism of network with a simulation module to recover the possible faults. KIRUBARAJAN, et al[14–15], developed a fault detection method using Markov models in large-scale systems. Petri-Net based alarm management is used to predict the hidden system faults[16]. HUANG et al[17], conducted the fault identification by comparing the signal signature with different network loads. HUANGSHUI, et al[18], presented a scanning based framework for online failure diagnosis, similar tool can be seen in Ref. [19]. However, in existing networked industrial systems, the information used in these algorithms is not generally available.

On the other hand, there are some works focused on IC induced fault detection and location identification. For example, a program called Manager was designed in the main network node for fault diagnosis by checking whether the sequence number in received message consistent with cache entry[20]. In industrial environment, time-based localization algorithm was developed for intermittent arc faults with fewer samples by measuring the three-phase voltages, currents, as well as cable parameters[21]. LEI, et al[22], developed a graph based fault diagnosis method. However, it was a knowledge-based method, and no quantitative measure was provided for decision-making. In addition, ZHAO, et al[23], established a statistical model to evaluate the distribution of error caused by IC, however, fault location identification was not addressed.

As can be seen from literature, there is no systematic fault location identification method for IC problem on CAN networks that could provide quantitative measures for fault location decision-making. Hence, the purpose of this study is to develop a new model based IC fault location identification method that is able to model the IC induced network errors using a stochastic approach to describe the patterns of errors and identify the location of the IC problem using measures that have physical and statistical meanings. The rest of the paper is organized as follows. We first define the researching problem in section 2. Then in section 3, the proposed method will be introduced in details, followed by experimental setup and case studies in section 4. Section 5 provides the conclusions and future work.

## 2 Problem Definition

As mentioned previously, an IC problem can deteriorate into a severe problem and result in serious system failure incidents if it is not treated properly. When an IC fault occurs, the data frames will be dropped and the error packets sent from the nodes in error active states will flood the bus. According to CAN specification[24], error packets sent from all of the nodes will composite an error frame, which consists of 6–12 consecutive dominant bits.

In this paper, we define the occurrence of error frame as an error event. As seen from Fig. 1, an error event is easy to detect, however, inferring the cause of error event is not straightforward since the information embedded in the error packets is limited, and the corrupted transmission data are discarded by CAN chips. Therefore, to locate the IC problem, the following two challenges must be addressed.

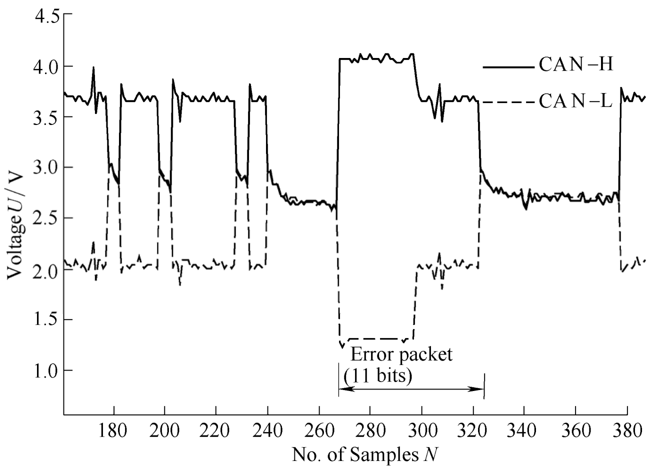



Fig. 1. Example of analog waveform on CAN bus when an error occurs

(1) How to automatically define and classify error events with limited information, and establish an appropriate model to describe the IC phenomenon whose parameters have physical and statistical meaning? It determines the quality and the quantity of the data to be collected, as well as the complexity of the fault location identification.

(2) How to systematically determine the location of the IC problem using the model parameters? Since it is of great importance to find the IC location accurately before the system is shut down.

To answer these questions, two assumptions are made: (1) there exists only one IC problem on the network and it is located on the drop cable of the network; and (2) the communication mode is polling or periodic.

It is noticed that some interferences, e.g. localized EMI problem, may show similar error phenomenon as IC problem. The isolation of these interferences from IC problem requires addition causality analysis, which is beyond the scope of this paper.

## 3 Proposed Methodology

In this paper, a model-based method is proposed for online fault location identification. The basic idea of the proposed method is that the location of the IC fault can be inferred by the patterns of the passively recorded network errors, which is described by the model of the faulty connected node. The overall procedure of the proposed method is illustrated in Fig. 2. Error event synchronizing is developed to recover the error event information upon each error. Then error event classification is performed based on the recovered information in each error for model selection. The most important step is error modeling, where node GZIP models are established using the corresponding error event sequences. Finally, IC location identification is conducted by evaluating the node GZIP model parameters. The details of procedure are introduced as follows.

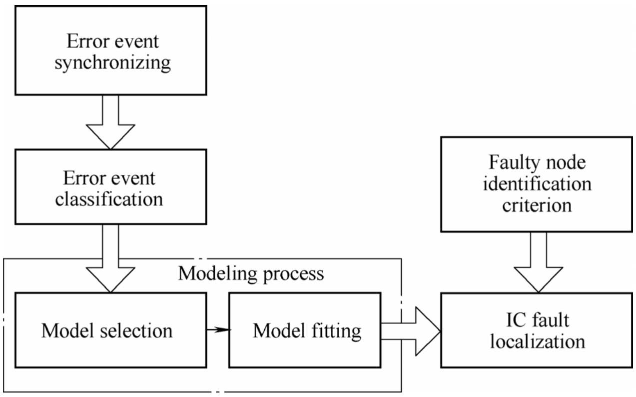



Fig. 2. Function block diagram of fault location identification

### 3.1 Error event synchronizing

Each error event is defined by a quadruple with the following elements: time stamp, the length of superimposed error frame, the interrupted frame source(the address of source node), as well as the source node that initiated the error. Since limited information is contained in the data link layer, physical layer error information must be collected to provide sufficient diagnostic information. Therefore, synchronizations of the error events are conducted where physical layer and data-link layer parameters are concurrently captured and integrated.

Fig. 3 briefly explains how to synchronize error events. From CAN specification, 6 consecutive dominant bits will constitute an error, which is used to trigger physical and data-link layer parameter acquisition. The timestamps of the errors are also recorded, as well as the length of superimposed error frames. As for the interrupted packet source and initiated error packet identification, recoverable address encoded into data packet is identified, or pattern recognition will be used to identify the destroyed data frame. The source of error frames are determined by comparing reference features with observed ones using analog features[25].

A sample of synchronized error event is illustrated in Table 1. This event can be described as follows: node 8 was transmitting messages when an error is detected at the 64.410452 s. Node 9 is the first node that responded the error and sent the error packet, followed by other nodes, thus eventually resulted in a 12-bits error frame.

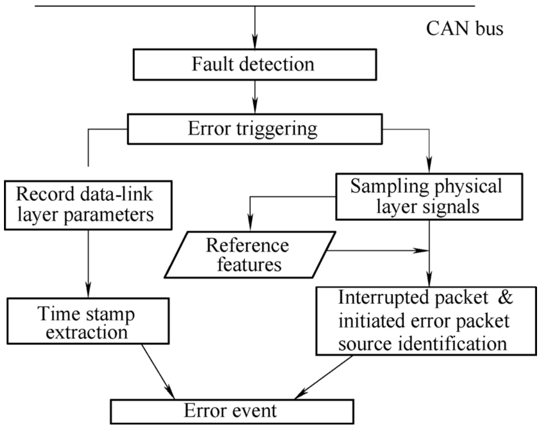



Fig. 3. Error event synchronizing

**Table 1. Illustration for synchronized error event on the segment from Fig. 1**

| IC detection | Time stamp $t$/s | Length of error frame $L$/bit | Analog waveform |
|---|---|---|---|
| Synchronized error event | 64.410 452 | 12 | P_8, E_9 |

P_8—Corrupted packet of node 8; E_9—Initiated error frame from node 9

### 3.2 Error event classification

In this subsection, we focused on which node an error event should be classified to, since statistic model needs to be established for each node through classified event sequences to reveal the relationship between network nodes and the error events. The classification procedure is based on the node status whether it is sending a data frame or initialing an error frame when the error event occurs.

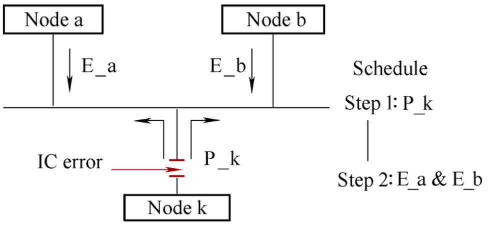
Fig. 4 illustrates two error scenarios when an IC fault occurs on the drop cable to Node k. “P_k” and “P_a” refer to packet sent from node “*k*” and node “*a*”, respectively. “E_k”, “E_a” and “E_b” stand for error packets from corresponding nodes. In scenario (a), node k is sending data frame when an IC problem occurs, all the other normal nodes will capture the IC induced packet logic error and respond at the same time. Hence, the compounded error frame is 6-bits long(identified as the node that provides maximum differential voltage). This error event is labeled as “T” pattern. In scenario (b), on the contrary, node error response is not simultaneous since IC can only affect the receiving result of node *k*. Therefore in this case node k will initiate the error packet, and the final length of error frame is 7–12 bits. Hence the error event is labeled as “R” pattern.

According to IC error generation mechanism, all the error events can be classified into aforementioned two patterns. As is seen in Fig. 4, in order to locate nodes with IC connection problem, it must significantly demonstrate both patterns. To do so, the following assumptions must be made.

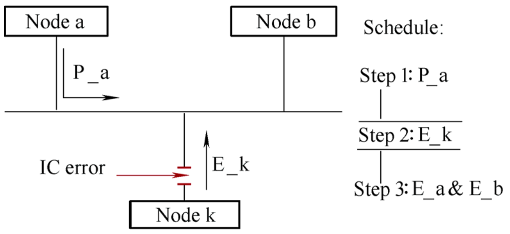
(1) Each error event from different nodes follows a separate distribution.

(2) The IC problem is consistent, so that meaningful distribution models can be established from the error

sequences.




(a) "T" pattern-derives when the faulty node is transmitting a message




(b) "R" pattern-derives when the faulty node is receiving a message

Fig. 4. Patterns of error events under IC problem

As seen from Fig. 4, the error events classified as "T" pattern can be generated only when the node with IC problem is occupying CAN bus the moment IC occurs, while "R" pattern error event can occur when the problematic node is in receiving mode. Hence, one can apply the essential diversity of the two patterns of error events to determine whether the node is experiencing IC problem. It is noted that in complex situations, node k in scenario (a) of Fig. 4 may occasionally send an error frame later than other nodes that response simultaneously, which results in a 7-12 bits error event. Although in this case the error cannot be distinguished from the error events generated in scenario (b), it will not affect the IC location identification analysis using the error events from scenario (a), which are usually sufficient to provide considerable evidence for IC problem.

As described previously, the classification of error events is conducted based on the patterns of error events. For instance, as illustrated in Table 2, a 12-bits error event emerges on bus at 5.891 752 s. The error frame is initiated by node 8, thus we classify this error event as R-type event by node 8. Later, another error event is observed at 6.887714 s, which is 6-bits long. It is generated at the moment node 8 is sending message. Thus we classify this error event as T-type event by node 8.

### 3.3 Modeling process

In section 3.2, for each node, the error events are classified as a timing sequence with different patterns. In order to provide statistical measures for fault location identification, a suitable stochastic model that reflects the actual distribution of error events is needed. Firstly, time partition window(T-Window) is used to divide the time sequence, thus transfer the time stamp sequences to counting processes of discrete events, which can be expressed as

$$N_T = [N_1(T), N_2(T), \cdots N_n(T)]^{\mathrm{T}}, \tag{1}$$

where $N_i(T)$ denotes the numbers of error event counted in the $i$th T-Window. The introduction of T-Window converts our modeling objective to the distribution of $N_T$.

**Table 2. Illustration for error events classified into corresponding network nodes**

| Node address | Time stamp $t$/s | Pattern of error event | Length of error Packet $L$/bit |
|---|---|---|---|
| 8 | 5.891 752 | R | 12 |
| | 6.887 714 | T | 6 |
| | 7.223 695 | T | 6 |
| | 7.247 693 | T | 6 |
| | ⋮ | ⋮ | ⋮ |
| 9 | 5.723 759 | R | 12 |
| | 6.623 810 | R | 9 |
| | 6.887 728 | R | 10 |
| | ⋮ | ⋮ | ⋮ |
| PLC | 10.763 559 | R | 12 |
| | ⋮ | ⋮ | ⋮ |
| | 64.722 427 | R | 8 |

Fig. 5 shows the overall modeling procedure. First, a proper model is selected to model the mixed type event sequences. Then parameter estimation is conducted, including initializing parameter, T-Window selection, recursive operation and checking for good-fitness of data distribution, in which an appreciate T-Window is very critical. The estimated parameters will be used for IC fault location identification.

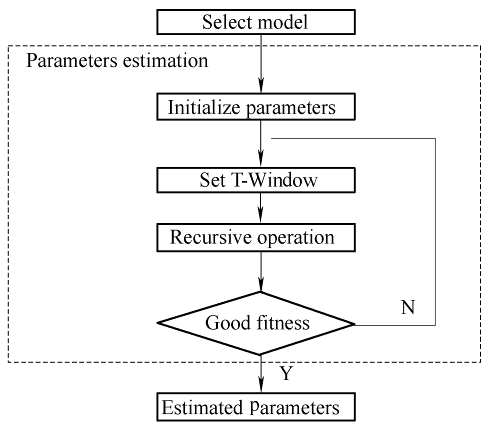



Fig. 5. Procedure of the GZIP model parameter estimation

#### 3.2.1 *Model selection*

In practice, the natural of IC problem determines that at majority of time, the CAN bus is normal and reliable. Hence there will be masses of zeros elements in $N_T$. Hence Zero-Inflated Poisson is often used to model the data[23]. Assuming that the event follows the Poisson distribution with parameter $\mu$ and the occurrence probability $P$, the probability mass distribution function that shapes the numbers of events within certain time, denoted as $X$, is described by

$$\begin{cases} P(X=0)=1-p+p\exp(-\mu), \\ P(X=x)=p\dfrac{\mu^x\exp(-\mu)}{x!}. \end{cases} \tag{2}$$

As discussed above, each network node may involve two patterns of the error events. Hence the generalized Zero-Inflated Poison(GZIP) model[26], extended on the basis of Zero-Inflated Poisson process, can be used if each kind of error event follows an independent distribution. Therefore the mixed type error event distributions can be expressed as follows:

$$P(X=x)=(1-\sum_{i=1}^{n}\lambda_i)\bullet I(X=0)+ \sum_{i=1}^{n}\lambda_i\frac{\mu_i^x\exp(-\mu_i)}{x!}(x\geqslant 0), \tag{3}$$

where $\lambda_i$ and $\mu_i$($i$=$t$ or $r$) are similar to parameters defined in Zero-Inflated Poisson model, while the index $t$ and $r$ refer to symbols of "T" pattern and "R" pattern of the error event, respectively. The indicator $I$ equals 1 if $X=1$, and equals to 0 otherwise. Hence every GZIP model consists of 4 parameters that determine the distribution of error events.

*3.2.2 Parameters estimation*

Expectation-Maximization method is used for point estimation of the GZIP model parameters by using the linear recursive expression[27]:

$$\begin{cases} \lambda_i^{k+1}=\dfrac{1}{n}\displaystyle\sum_{q=1}^{N}\Pr(i\mid X_q,\theta^k), \\ \mu_i^{k+1}=\dfrac{\displaystyle\sum_{q=1}^{N}X_q\bullet\Pr(i\mid X_q,\theta^k)}{\displaystyle\sum_{q=1}^{N}\Pr(i\mid X_q,\theta^k)}, \end{cases} \quad i=t\text{ or }r, \tag{4}$$

where $N$ represents the numbers of elements of $N_T$, and $\Pr(i\mid X_q,\theta^k)$ means the posterior probability that $X_q$ comes from the $i$th error event given the presently iterated parameters $\theta^k$ calculated as follows:

$$\Pr(i\mid X_q,\theta^k)=\frac{\lambda_i^k P_i(X_q\mid\mu_i^k)}{\displaystyle\sum_{j=0}^{n}\lambda_j^k P_j(X_q\mid\mu_j^k)}. \tag{5}$$

Eq. (4) provides an effective process for parameter iterations. However, the initial values of GZIP model parameters is critical to the convergence of the estimation. Let us denote $t_{\max}$ as the time of the last sample, $N_i$ and $N_{\text{tol}}$ stand for the numbers of the $i$th error event and total number of the error events, respectively. The parameters in GZIP model can be initialized as follows:

$$\begin{cases} \overline{\lambda}_i=(1-\lambda_0)N_i/N_{\text{tol}}, \\ \overline{\mu}_i=N_i\bullet T/t_{\max}, \end{cases} \quad i=t\text{ or }r. \tag{6}$$

Eq. (6) provides a symmetrical initialization approach based on the ratio of zeros $\lambda_0$.

In addition, when calculating the probability of each pattern of error events, the width of T-Window($T$) must be selected appropriately. several criteria are imposed as follows.

(1) Statistical meaning should be guaranteed, namely,

$$\lambda_i,\mu_i\geqslant 0\text{ and }\sum_i\lambda_i\equiv 1,\ \ i=0,t\text{ or }r. \tag{7}$$

(2) The fitted GZIP model musts consist of zeros. That is, $\lambda_0>0$.

(3) The fitted model should reflect the diversity of different error events. Hence $\chi^2$ test is used to minimizing the difference between fitted model and empirical data. In $\chi^2$ test, if a model has $k$ (e.g., $k$=4 in GZIP model) parameters, then the class of data is at least $k$+1. In our application, let us denote $f_G(\cdot)$ and $f_E(\cdot)$ as the distribution density function of the GZIP model and empirical data, respectively, and the optimal T-Window can be obtained by

$$T=\underset{n\geqslant 5}{\operatorname{argmin}}\sqrt{\sum_{j=0}^{n}\left|f_G\left(N_j(T)\right)-f_E\left(N_j(T)\right)\right|^2\Big/n}. \tag{8}$$

Since solving Eq. (8) requires Eq. (6), iterative parameter estimation procedure for GZIP model is developed. For further discussion, we denote the estimated parameters as $\hat{\theta}\left(\hat{\theta}:\lambda_t,\mu_t,\lambda_r,\mu_r\right)$.

The likelihood confidence region estimation is applied to calculate the confidence intervals of parameters in GZIP model. The confidence region $\theta'$ with significance level $\alpha$ is

$$\left\{\theta':\ 2\ln\frac{\prod_{i=1}^{n}f_G\left(N_i(T);\hat{\theta}\right)}{\prod_{i=1}^{n}f_G\left(N_i(T);\theta'\right)}\leqslant\chi^2_{1-\alpha}(k)\right\}= \left\{\theta':\prod_{i=1}^{n}f_G\left(N_i(T);\theta'\right)\geqslant \exp(-\chi^2_{1-\alpha}(k)/2)\prod_{i=1}^{n}f_G\left(N_i(T);\hat{\theta}\right)\right\}. \tag{9}$$

Because of the complicated structure in Eq. (9) with multiple parameters $\theta' = (\lambda_t', \mu_t', \lambda_r', \mu_r')$, a standardization parameter exploration procedure is applied to handle two parameter changes: single $\lambda_i'(i = t \text{ or } r)$ changes and single $\mu_i'$ changes. Suppose $\lambda_i$ changes to $\lambda_i + \Delta\lambda$, $\lambda_k$ $(k \neq i, k \in \{t, r\})$ also needs to be changed to satisfy $\sum_i \lambda_i \equiv 1$, and hence the corresponding strategy is

$$\begin{cases} \lambda_i' = \dfrac{\lambda_i + \Delta\lambda}{1 + \Delta\lambda}, \\ \lambda_k' = \dfrac{\lambda_k}{1 + \Delta\lambda}. \end{cases} \tag{10}$$

Compared with the changes on $\lambda_i$, single shift on $\mu_i$ does not affect other parameters.

### 3.4 IC fault location identification

Fault location identification is conducted by interpreting the confident intervals of the parameters of the GZIP models. As described previously, "T" pattern error events occurs only when the faulty node is sending packets, such 6-bits long error events could not occur to a normal node. Therefore we focused on the confidence interval of parameters of the "T" pattern error events $I_t$. Therefore the location of IC is can be determined if the "T" pattern error events of the node is significant statistically. Fig. 6 shows the procedure for IC location identification. A node is labeled to have IC problem only if its corresponding GZIP model parameter $\lambda_t$ is statistically significant, that is, zero is not included in the confident region $I_t$. Otherwise this node is normal with 1-$\alpha$ confidence level. In the case that two or more IC problems exist concurrently, the fault location identification procedure remains unchanged.

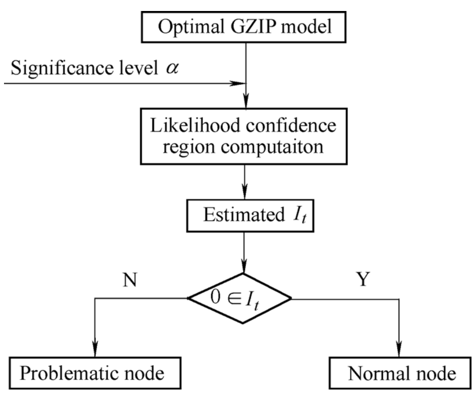



Fig. 6. Evaluation criterion for identifying problematic node

## 4 Case Study

### 4.1 Test-bed setup

To demonstrate the proposed fault location identification method, laboratory experiments have been conducted. A 3-node experimental CAN network system is illustrated in Fig. 7(It should be noted that the proposed method is general to more nodes.) The network is set to communicate at $5\times10^5$ bit/s using polling mode with 10 ms polling intervals controlled by a PLC scanner module. The IC problem is emulated on a drop cable through a high-speed on-off switch controlled by a computer, and the switching interval($\Delta t$) between two adjacent disconnections event follows a Poisson distribution $p = \mu^{\Delta t} \exp(-\mu)/\Delta t!$. In this case study, the average switching interval is set to 75 ms. The duration of each disconnection is half bit width. Moreover, an FPGA based error detector has been developed to provide trigger signal to multi-layer data acquisition system (DAQ) upon each error. The constructed test-bed is shown in Fig. 8.

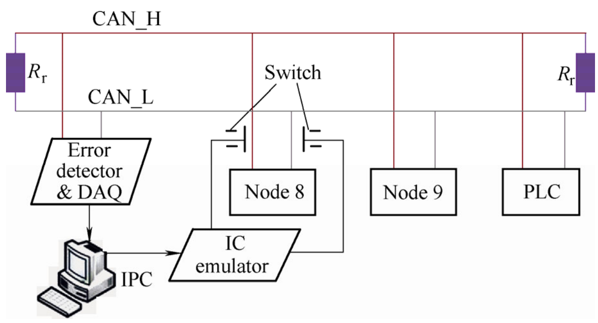



Fig. 7. Schematic of experiment design for IC

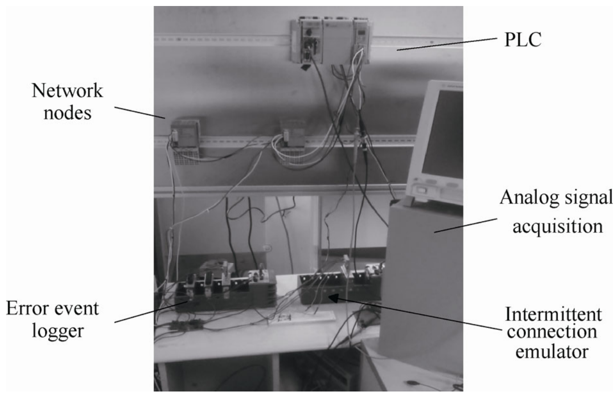



Fig. 8. Layout of test-bed

### 4.2 Result analysis

#### 4.2.1 *Model evaluation*

Since each node will show different node-error interactions under IC problem, GZIP models for each node are constructed. Fig. 9 illustrates the modeling result of node 8. As seen from the figure, the fitting distribution with the expected model parameters

$$\hat{\theta}\left(\hat{\theta} : [\lambda_t, \mu_t, \lambda_r, \mu_r] = [0.323\,9, 1.447\,5, 0.620\,7, 3.657\,9]\right)$$

is very close to the empirical data histogram. Furthermore, the peak value of probability density is reached when the number of error events within the T-Window is between $\mu_t$ and $\mu_r$, which indicates that the two patterns of error events has been sufficiently shaped by these two sets of parameters together while "T" pattern error events tend to play a dominant role.

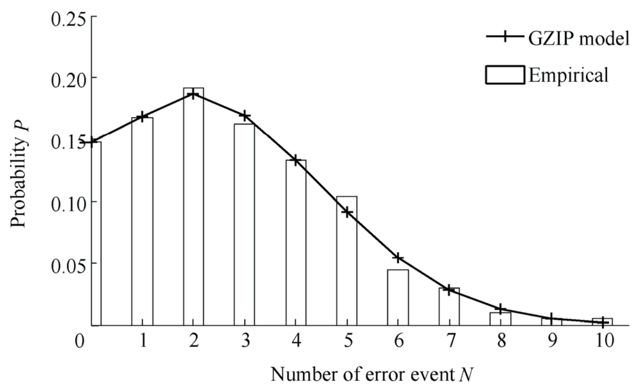



Fig. 9. GZIP modeling result for node 8

For comparison, the parameters of GZIP model for each single node are shown in Table 3. As it can be seen in the table, $\lambda_t$ and $\mu_t$ of node 9 and PLC all equals to 0, thus their GZIP models degraded into Zero-Inflated Poisson distribution, since no “T” pattern error events exists. Reasonable interpretation can be made from Fig. 10 where the characteristic of Zero-Inflated Poisson for node 9 have been revealed, that is, the probability density decreases sharply against the increasing numbers of error events, since the distribution is represented by “R” pattern error events with Poisson parameter $\mu_t$=2.572 9. Moreover, since node 9 provides maximum differential signal voltage, it is identified as the source of initiated error packets when several nodes send error packet simultaneously. Hence a local maximum is observed in Fig. 10. In addition, the GZIP modeling result for the error event for PLC is shown in Fig. 11.

**Table 3. Parameters of GZIP model of each node**

| Property | Node 8 | Node 9 | PLC |
|---|---|---|---|
| T-Window $t$/s | 0.8160 | 0.8310 | 4.1620 |
| $\lambda_t$ | 0.3239 | 0 | 0 |
| $\mu_t$ | 1.4475 | 0 | 0 |
| $\lambda_r$ | 0.6207 | 0.7057 | 0.7090 |
| $\mu_r$ | 3.6579 | 2.5729 | 1.3264 |

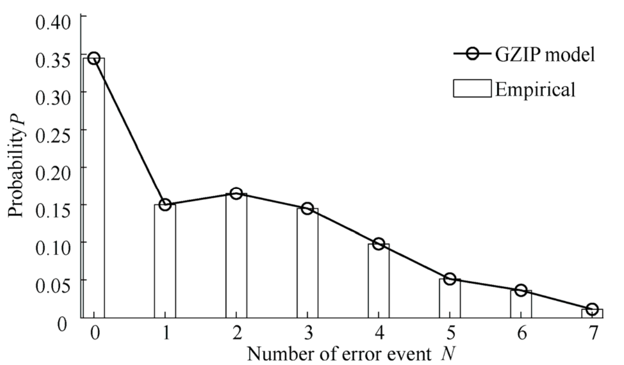



Fig. 10. GZIP modeling result for node 9

The estimated parameters in Table 3 show a preliminary answer to the IC location identification problem, where the GZIP model for node 8 shows two patterns of error events. Hence, node 8 can be preliminary identified to have IC problem. However, further analysis needs to be conducted by determining whether the result is statistically sound.

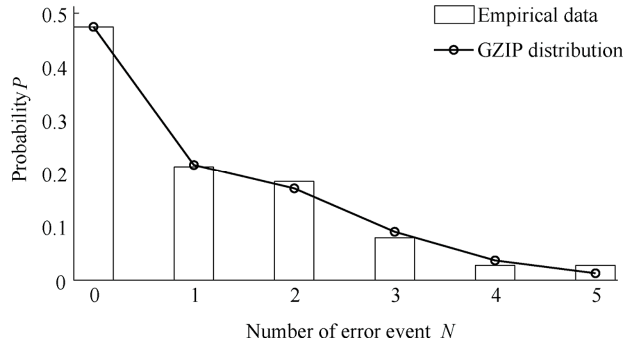


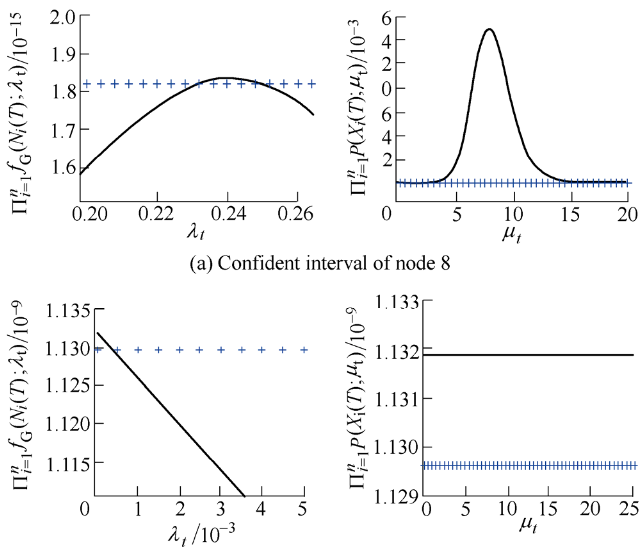

Fig. 11. GZIP modeling result for PLC

### 4.2.2 *Final fault location identification*

Given significance level $\alpha$=0.05, the confidence intervals $I_t$ of GZIP model parameters for each node are listed in Table 4. As it can be seen from the table, every parameter interval includes its estimated value in its corresponding GZIP models, which demonstrates the effectiveness of the proposed confidence region computation method. Detailed calculations for the parameters are illustrated in Fig. 12.

**Table 4. Parameter confidence intervals of each node**

| Node | $\lambda_t$ | | $\mu_t$ | |
|---|---|---|---|---|
| address | Lower limit | Upper limit | Lower limit | Upper limit |
| 8 | 0.282 0 | 0.328 5 | 1.424 6 | 16.475 7 |
| 9 | 0 | $3.9\times10^{-4}$ | 0 | $+\infty$ |
| PLC | 0 | $5.1\times10^{-4}$ | 0 | $+\infty$ |




(a) Confident interval of node 8

(b) Confident interval of node 9

(Confident intervals from PLC are similar to node 9)

Fig. 12. Likelihood confident intervals of model parameters for the network nodes

(The likelihood functions are denoted by line with "+" maker)

According to the criterion of fault location identification procedure in Fig. 6, we can determine that node 8 is the problematic node with IC problem, since $\lambda_t$ of node 8 is significant. On the contrary, $I_t$ from both node 9 and PLC include 0. In addition, the upper limits of $\lambda_t$ from node 9 and PLC are all sufficiently small, which means no "T" pattern error events are possible to occur on these two nodes. Therefore, with 962 errors collected in this study, the 6-bits long error events observed on CAN bus are totally from node 8 associated with random IC faults, which agrees well with the experimental setup where IC problem is set on the drop cable to node 8. Additional analysis shows that $I_t$ of normal nodes would remain stable when the sample size increases, which ensures the robustness of fault location identification method proposed.

## 5 Conclusions

(1) A novel model based IC location identification method is developed for IC fault detection and isolation on CAN network using error event sequence information.

(2) A GZIP model based IC location identification algorithm is developed to determine whether a network node is problematic or not by testing the parameters of the GZIP model using the confident intervals of the parameters.

(3) Experiments are conducted to illustrate the proposed method on a 3-node CAN network testbed. Experimental results show that when the IC problem is located on a drop cable, the parameters of GZIP model for the corresponding connecting node are statistically significant.

Future work will be focused on the complex IC situations where multiple IC problems on different locations, including both drop cable and backbone of the network.

## References

[1] FARSI M. An overview of controller area network[J]. *Computing and Control Engineering Journal*, 1999, 10(3): 113–120.

[2] PRASAD V B. Markovian model for the evaluation of reliability of computer networks with intermittent faults[C]//*IEEE International Symposium on Circuits and Systems*, Singapore, Jun 11–14, 1991: 2084–2087.

[3] LIAN F L, MOYNE J, TILBURY D. Performance evaluation of control network[J]. *IEEE Control Systems Magazine*, 2001, 21(1): 66–83.

[4] TRAN E. *Multi-bit error vulnerabilities in the controller area network protocol*[M]. Pittsburgh: Carnegie Mellon University Press, 1999.

[5] ZHAO Z G. A hierarchy management framework for automated network fault identification[C]//*IEEE International Conference on Wireless Communications, Networking and Mobile Computing*, Dalian, China, Oct 12–17, 2008: 1–4.

[6] RUFINO J, VENSSIMO C A P, ARROZ G. Enforcing dependability and timeliness in controller area networks[C]//*IEEE Industrial Electronics, IECON 32nd Annual Conference*, Paris, France, Nov 6-10, 2006: 3755–3760.

[7] HANSSON H A, NOLTE T, NORSTROM C, et al. Integrating reliability and timing analysis of CAN-based systems[J]. *IEEE transactions on Industrial Electronics*, 2002, 49(6): 1240–1250.

[8] KIM W, JI K, AMBIKE A. Networked real-time control strategy dealing with stochastic time delays and packet losses[J]. *Journal of Dynamic Systems, Measurement, and Control*, 2006, 128(3): 681–685.

[9] AYSAN H, THEKKILAKATTIL A, PUNNEKKAT S. Efficient fault tolerant scheduling on controller area network CAN[C]//*Emerging Technologies and Factory Automation(ETFA)*, Bilbao, Spain, Sep 13–16, 2010: 1–8.

[10] AYSAN H, PUNNEKKAT S. Fault tolerant scheduling on controller area network(CAN)[C]//*2010 13th IEEE International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing Workshops*, Carmona, Spain, May 4–7, 2010: 226–232.

[11] BHATTACHARYYA S, CHANDRA V, KUMAR R. A discrete event systems approach to network fault management: detection & diagnosis of faults[C]//*Proceeding of the 2004 American Control Conference*, Boston Sheraton Hotel, Boston, Massachusetts, Jun 30–Jul 2, 2004: 5108–5113.

[12] JIANG S, KUMAR R, GARCIA H E. Diagnosis of repeated/ intermittent failures in discrete event systems[J]. *IEEE Transactions on Robotics and Automation*, April, 2003, 19(2): 310–313.

[13] ANGSKUN T, BOSILCA G, FAGG G, et al. Reliability analysis of self-healing network using discrete-event simulation[C]//*Seventh IEEE International Symposium on Cluster Computing and the Grid*, Rio de Janeiro, Brazil, May 14–17, 2007: 437–444.

[14] KIRUBARAJAN T, MALEPATI V N, DEB S, et al. Fault detection algorithms for real-time diagnosis in large-scale systems[C]//*Component and Systems Diagnostics, Prognosis, and Health Management*, Calcutta, India, Dec 18–20, 2001,4389: 243–254.

[15] XU Y, GE M, DU R. A real-time monitoring and diagnosis systems for manufacturing automation[C]//*International Conference on Robotics & Automation*, New Orleans, United states, Apr 26–May 1, 2004, 2: 1424–1429.

[16] CHAO C S, LIU A C. An alarm management framework for automated network fault identification[J]. *Computer Communications*, 2004, 27(13): 1341–1353.

[17] HUANG Y, CHENG W T. Intermittent scan chain fault diagnosis based on signal probability analysis[C]//*Design, Automation and Test in Europe Conference and Exhibition*, Paris, France, Feb 16–20, 2004, 2: 21072.

[18] HUANGSHUI H, GUIHE Q. Online fault diagnosis for controller area networks[C]//*Fourth International Conference on Intelligent Computation Technology and Automation*, Shenzhen, Guangdong, China, Mar 28–29, 2011,1: 452–455.

[19] WINSTEAD V, KOLMANOVSKY I. Observers for fault detection In networked systems with random delays[C]//*Proceeding of the 2004 American Control Conference*, Boston, United states, Jun 30–Jul 2, 2004: 2457–2462.

[20] HUANGSHUI H, GUIHE Q. Online fault diagnosis for controller area networks[C]//*2011 Fourth International Conference on Intelligent Computation Technology and Automation*, Shenzhen, Guangdong, China, Mar 28–29, 2011, 1: 452–455.

[21] ALAMUTI M M, NOURI H, TERZIJA V. Intermittent fault location in distribution feeders[J]. *IEEE Transactions on Power Delivery*, Jan, 2012, 27(1): 96–103.

[22] LEI Y, DJURDJANOVIC D. Diagnosis of intermittent connections for DeviceNet[J]. *Chinese Journal of Mechanical Engineering*, 2010, 23(5): 606–612.

[23] ZHAO J, LEI Y. Modeling for early fault detection of intermittent connections on controller area networks[C]//*IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, Kaohsiung, Taiwan, China, Jul 11–14, 2012: 1135–1140.

[24] BOSCH. CAN specification version 2.0[J]. *Bosch*, Sep, 1991.

[25] LEI Y, DJURDJANOVIC D, BARAJAS L, et al. DeviceNet network health monitoring using physical layer parameters[J]. *Journal of Intelligent Manufacturing*, 2011, 22(2): 289–299.

[26] CHEN N, ZHOU S, CHANG T, et al. Attribute control charts using generalized zero-inflated Poisson distribution[J]. *Quality and Reliability Engineering International*, 2008, 24(7): 793–806.

[27] LIU Z, ALMHANA J, CHOULAKIAN V. Recursive EM algorithm for finite mixture models with application to internet traffic

modeling[C]//*IEEE Second Annual Conference on Communication Networks and Services Research*, Fredericton, Canada, May 19–21, 2004: 198–207.

## Biographical notes

LEI Yong, born in 1976, is an associate professor at *State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, China*. He received his BS degree in control science and engineering from *Huazhong University of Science and Technology, China*, his MS degree in manufacturing and automation from *Tsinghua University, China*, and his PhD degree in mechanical engineering from *University of Michigan, Ann Arbor, USA*. His research interests include monitoring and fault diagnosis of the networked automation systems, statistical quality control, and surgical robots.
E-mail: ylei@zju.edu.cn

YUAN Yong, born in 1989, is a master candidate at *State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, China*. He received his BS degree in mechanical engineering from *Hunan University, China*. His research interests include monitoring and fault diagnosis of the networked automation systems and applied statistics.
E-mail: 21125110@zju.edu.cn

SUN Yichao, born in 1990, is a master candidate at *State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, China*. He received his BS degree in mechanical design, manufacturing and automation from *East China University of Science and Technology, China*. His research interests include fault diagnosis and robust control of the networked automation systems.
E-mail: 21325082@zju.edu.cn